

Mayer-Cluster Expansion of Instanton Partition Functions and Thermodynamic Bethe Ansatz

Carlo Meneghelli ^{a,b} and Gang Yang ^c¹

^a *Fachbereich Mathematik, Universität Hamburg
 Bundesstraße 55, 20146 Hamburg*

^b *Theory Group, DESY
 Notkestraße 85, 22603 Hamburg, Germany*

^c *Institut für Physik, Humboldt-Universität zu Berlin
 Newtonstraße 15, 12489 Berlin, Germany*

Abstract

In [18] Nekrasov and Shatashvili pointed out that the $\mathcal{N} = 2$ instanton partition function in a special limit of the Ω -deformation parameters is characterized by certain thermodynamic Bethe ansatz (TBA) like equations. In this work we present an explicit derivation of this fact as well as generalizations to quiver gauge theories. To do so we combine various techniques like the iterated Mayer expansion, the method of expansion by regions, and the path integral tricks for non-perturbative summation. The TBA equations derived entirely within gauge theory have been proposed to encode the spectrum of a large class of quantum integrable systems. We hope that the derivation presented in this paper elucidates further this completely new point of view on the origin, as well as on the structure, of TBA equations in integrable models.

¹ carlo.meneghelli@desy.de, gang.yang@physik.hu-berlin.de

Contents

1	Introduction	3
1.1	NS's correspondence	7
2	Mayer-Cluster expansion	8
2.1	Only long range interactions	9
2.2	Only short range interactions	11
2.2.1	Dilogarithm from a sum over residues	12
2.2.2	By Mayer expansion and separation of regions	13
2.3	Expansion of the instanton partition function	16
2.3.1	Iterated Mayer expansion	16
2.3.2	Iterated Mayer expansion for the Nekrasov partition function	17
3	Derivation of the TBA	19
3.1	Perturbative expansion	19
3.2	TBA as saddle-point of a $(0+1)$ -dimensional path integral	20
3.2.1	Multi-particle interactions	22
4	Generalization to quiver gauge theories	23
4.1	TBA for quiver gauge theories	25
5	Conclusion and Discussion	26
A	On the integral representation form of the Nekrasov partition function	28
A.1	D-instanton action and its localization	29
A.2	The instanton partition function from 5d perspective	30
A.3	Building blocks for quiver gauge theories	32
A.4	Contour and classification of poles	33
B	Useful formulas	34
B.1	Computing (2.16)	34
B.2	Exponentiation of residues	34
B.3	Illustrative Examples for Iterated Mayer expansion	35
C	Expansion by regions for instanton partition functions	35
C.1	Checks of expansion by regions	38
D	Alternative derivation of (2.43)	41

1 Introduction

A number of remarkable connections have been observed between gauge theories and integrable systems. They appear to be useful to increase our understanding of both subjects. On the one hand, using powerful integrability techniques one may hope to solve certain gauge theories non-perturbatively. On the other hand, gauge theory can help to formulate and solve integrable system. A spectacular example is given by planar $\mathcal{N} = 4$ super Yang-Mills theory, for which in the last ten years or so tremendous progress has been achieved in solving the theory based on the underlying integrability and on the AdS/CFT correspondence, see [1] for a recent review. There exists another amazing connection between gauge theories and integrable models. The gauge theories in this case are $\mathcal{N} = 2$ supersymmetric gauge theories. They do not need to be planar but the connection with integrable models is restricted to a special class of supersymmetric observables. In this paper we focus on an important object in this class, the so-called instanton partition function and its relation with quantum integrable systems.

In the groundbreaking work of Seiberg and Witten [2, 3], the exact solution for the low energy effective action of certain $\mathcal{N} = 2$ gauge theories was proposed based on holomorphicity properties and electromagnetic duality. The low energy dynamics are encoded in a single object, called prepotential $\mathcal{F}(\vec{a})$. It is a holomorphic function on the Coulomb moduli space, with coordinates \vec{a} , and can be reconstructed from the so-called Seiberg-Witten (SW) curve and SW differential. Shortly after, it was realized that this description provides a direct connection between $\mathcal{N} = 2$ gauge theories and classical algebraically integrable systems [4, 5, 6], see e.g. [7] for a pedagogical introduction. These are essentially a complex analogue of integrable systems in the sense of Liouville.

The challenging program of obtaining the Seiberg-Witten prepotential by a direct gauge theory calculation, developed on [8, 9, 10], was finalized in [11]. This problem was solved using powerful localization techniques. Interestingly, this calculation produced a two-parameter, called ϵ_1 and ϵ_2 , deformation of the prepotential. The SW prepotential can be obtained as

$$\mathcal{F}(\vec{a}; q) = \lim_{\epsilon_1, \epsilon_2 \rightarrow 0} \epsilon_1 \epsilon_2 \log \mathcal{Z}(\vec{a}, \epsilon_1, \epsilon_2; q), \quad (1.1)$$

where we add the explicit dependence on the coupling constant q , but suppress dependence on further parameters such as masses. The partition function \mathcal{Z} receives tree level, one loop and instanton contributions. The latter part is usually referred to as Nekrasov instanton partition function.

The parameters ϵ_1, ϵ_2 correspond to a Lorentz rotation Ω , thus the name Ω -deformation, that encodes certain twisted boundary conditions for the four dimensional gauge theory. It was first introduced in [8, 9] in order to regularize the volume of the instanton moduli space. The Ω -deformation can be understood in a simple way by considering the five dimensional lift of the $\mathcal{N} = 2$ theory, further compactified on a circle [12]. In this set-up it is interpreted as twisting of \mathbb{R}^4 by a Lorentz rotation while going around the circle. Introducing

a two-parameter generalization of the prepotential triggered a huge progress. An important example is the connection with topological strings. Upon taking $\epsilon_1 = -\epsilon_2 = g_s$, the gauge theory partition function reproduces the topological string partition function with g_s as genus parameter [13, 14]. The question of what is the topological string theory analog of the $\epsilon_1 \neq \epsilon_2$ case leads to the definition of so-called “refined” topological strings [15, 16]. Another spectacular example of progress driven by the calculation of the Nekrasov partition function is given by the Alday-Gaiotto-Tachikawa (AGT) correspondence [17].

More recently, Nekrasov and Shatashvili [18] proposed that upon taking the limit $\epsilon_2 \rightarrow 0$ and interpreting ϵ_1 as Plank constant, one obtains a correspondence between supersymmetric vacua of a given gauge theory and eigenstates of the corresponding *quantum* integrable model. The relation between the SW prepotential and classical integrable systems is thus quantized. This is usually called Nekrasov-Shatashvili (NS) limit and will be the main focus of this paper. The central role in this correspondence is played by the so-called twisted superpotential $\mathcal{W}(\vec{a}, \epsilon_1; q)$ defined as

$$\mathcal{W}(\vec{a}, \epsilon_1; q) = \lim_{\epsilon_2 \rightarrow 0} \epsilon_2 \log \mathcal{Z}(\vec{a}, \epsilon_1, \epsilon_2; q). \quad (1.2)$$

The Nekrasov-Shatashvili proposal is that, once this function is known, the eigenstates of the quantum integrable system are classified by solutions of the following quantization condition

$$\exp\left(\frac{\partial \mathcal{W}(\vec{a})}{\partial a_\lambda}\right) = 1, \quad \lambda = 1, \dots, r, \quad (1.3)$$

where r is the rank of the gauge group. These equations identify the twisted superpotential \mathcal{W} with the so-called Yang-Yang (YY) function [19] of the quantum system. The proposed correspondence provides an efficient general mechanism to define and solve quantum integral models. Remarkably, it can be argued that the instanton part of the prepotential $\mathcal{W}_{\text{inst}}$, defined via (1.2), can be characterized as the solution of certain non-linear integral equation of the Thermodynamic Bethe Ansatz (TBA) type [20]. The main goal of this paper is to develope some of the ideas presented in [18] to give a more explicit derivation, as well as some generalizations, of such TBA equations.

The proposal above originates as some sort of extension of the so-called Bethe/gauge correspondence [21, 22]. The latter is based on the observation that the vacuum equations of two dimensional $\mathcal{N} = 4$ gauge theories, broken to $\mathcal{N} = 2$ by twisted masses, coincides with Bethe equations for integrable models. The two dimensional twisted superpotential is equal to the YY function that encodes the Bethe equations and Coulomb parameters correspond to the Bethe roots. The generators of chiral ring of the gauge theory [23] are mapped to Hamiltonians of the integrable model, while their expectation values mapped to the corresponding eigenvalues. In this way one obtaines a large class of integrable models whose spectrum is characterized by traditional, possibly nested, Bethe equations. Many integrable models do not belong to this class. The simplest example is given by the quantum Toda chain, see e.g. [24], whose classical limit is connected to four dimensional pure $SU(N)$ SYM. From the insight of the Bethe/gauge correspondence it is then natural to propose [18]

the two dimensional twisted superpotential which corresponds to such integrable models. It is the effective low energy action for the four dimensional $\mathcal{N} = 2$ gauge theory subject to an Ω -background that preserves two dimensional $\mathcal{N} = 2$ super-Poincaré symmetry, namely $\epsilon_2 = 0$. This observation provides a strong motivation for the above correspondence. An interpretation of this correspondence was given using brane constructions in [25]. A further essential step in understanding the nature of the relation between quantum integrable systems and gauge theories has been presented in [26].

The proposal of Nekrasov and Shatashvili has inspired many other studies. Let us mention a few. In [27, 28] it was shown that, similarly to the prepotential \mathcal{F} , the twisted superpotential $\mathcal{W}(\epsilon_1)$ can be obtained by calculating period integrals of a suitably deformed SW differential. This analysis is also inspired by the AGT correspondence, by which the NS limit corresponds to the semiclassical limit of Liouville CFT [29, 30].

In our work the Coulomb parameters will be assumed to be in generic positions. Extra considerations are needed if they take special values. For example in the conformal $SU(N)$ SYM with $N_f = 2N$, if the Coulomb parameters a are set to be equal to $m_{\text{fund}} - n\epsilon_1$, where $n \in \mathbb{Z}^N$, one can quantize the corresponding integrable system to obtain a length N spin-chain with infinite dimensional highest weight representations of \mathfrak{sl}_2 at each site [31, 32, 33]. Its spectrum is described in terms of traditional Bethe Ansatz equations. Such developments triggered the discovery of a number of new dualities between various integrable models [34, 35, 36, 37, 38]. The NS proposal has also inspired various studies in (refined topological) string theories [39, 40, 41, 42, 43, 44, 45, 46] where the the general Ω -background plays a crucial role.

Despite the importance of this correspondence, the precise mechanism by which the instanton part of the twisted superpotential defined in (1.2) turns out to be characterized as the solution of TBA equations is still to be elucidated. In this paper, we will fill this gap. In order to fully prove the NS's proposal, at least in some example, one should be able to show that the same TBA equation characterizes the spectrum of the corresponding integrable model. In the case of pure $SU(N)$ SYM, which corresponds to the periodic Toda chain with N sites, this was achieved in [24]. This interesting problem will be studied elsewhere [47]. In the following we briefly outline the main ingredients used in our derivation of the TBA equations for (1.2), as well as the structure of the paper.

As pointed out in [18], it is convenient to start with the contour integral form of the instanton partition function. In this representation the instanton partition function can be interpreted as the partition function of a non-ideal gas of particles. The particular structure of the two-particle interaction potential makes the study of the $\epsilon_2 \rightarrow 0$ limit rather subtle. More precisely, this potential is the sum of a short-range (of order ϵ_2) strongly attractive piece and a long-range interaction part. In Sections 2.1 and 2.2, we consider the simplified situation in which either the long- or short-range part is set to zero. In order to study these simplified partition functions in the $\epsilon_2 \rightarrow 0$ limit, we combine a number of techniques like Mayer expansion [48] (a standard method in statistical mechanics) and the method

of expansion by regions [49] (a powerful method to compute Feynman integrals in small parameter expansions). For the case with only long-range interactions, the $\epsilon_2 \rightarrow 0$ limit turns the logarithm of the partition function into a sum over certain tree graphs. On the other hand, the free energy corresponding to only short-range interactions gives rise to the dilogarithm function Li_2 , which can be shown either by direct residue calculation of relevant integrals or via Mayer expansion together with the method of expansion by regions.

In order to study the full partition function we find it convenient to use an *iterated* version of Mayer expansion, see [50]. This expansion effectively creates a new partition function whose “new particles” are *clusters* of the original particles. The interaction within each cluster is governed by the short-range interaction, the one between different clusters by the long-range part. The iterated Mayer expansion thus produces an expression for the twisted superpotential $\mathcal{W}_{\text{inst}}$ as a sum over tree graphs, with vertices given by *clusters*. This expansion is carried over in some details in Section 2.3. The expression can be compared to high order in the instanton number with the expression coming from the solution to the TBA equation as discussed in Section 3.1, providing direct non-trivial check of the equality.

There is an elegant way to prove that this equality holds to all orders in the instanton counting parameter q . It is based on rewriting the grand canonical partition function of the non-ideal gas in terms of a $(0+1)$ -dimensional path integral. The analysis needs some special care as the potential has an unusual feature of depending in a singular way on \hbar , which is identified with ϵ_2 in this case. A slight modification of the argument in [51], together with the calculation of the contribution from the short range interactions corresponding to the dilogarithm, shows that the instanton partition function in the $\epsilon_2 \rightarrow 0$ limit is obtained by the saddle point evaluation of the path integral. The saddle point equations are nothing but the TBA equations. This is explained in details in Section 3.2.

In Section 4 we present a generalization of the TBA equations corresponding to quiver gauge theories. More precisely, we consider quivers characterized by a Dynkin diagram of ADE , or $\widehat{A}\widehat{D}\widehat{E}$ type. The twisted superpotential for such theories is shown to satisfy a set of coupled TBA equations with one equation for each node of the quiver and couplings corresponding to edges in the quiver. The derivation is a simple extension of the one for the single gauge group case. This is so as the short range interaction, responsible for the clustering of particle, is non-vanishing only for particles corresponding to the same gauge group factor in the quiver.

In order not to overload the main text, in Appendices we include some review material together with a few technical points concerning the derivation. A review of the contour integral form of the instanton partition function is given in Appendix A. Some useful formulas are collected in Appendix B. A discussion of the method of expansion by regions is given in Appendix C. In Appendix D, we present an alternative derivation of a tree graphs expansion of the instanton partition function.

The full partition function is a product of three terms $\mathcal{Z} = \mathcal{Z}_{\text{tree}} \mathcal{Z}_{\text{1-loop}} \mathcal{Z}_{\text{inst}}$. In this paper we will be only concerned with the study of the instanton part $\mathcal{Z}_{\text{inst}}$. For this reason

from now on it will be simply denoted by \mathcal{Z} .

1.1 NS's correspondence

To complete the introduction, we present the integral representation of the instanton partition function for pure $SU(N)$ $\mathcal{N} = 2$ super-Yang-Mills and the corresponding TBA equation. The derivation of the TBA starting from the gauge theory expression of the instanton partition function is the main goal of the paper.

Instanton partition function

The instanton partition function can be written in a contour integral representation as

$$\mathcal{Z} = \sum_{k=0}^{\infty} \left(\frac{\epsilon}{\epsilon_1 \epsilon_2} \right)^k \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} Q(\phi_I) \prod_{J>I}^k \mathcal{D}(\phi_{IJ}), \quad (1.4)$$

where $\phi_{IJ} = \phi_I - \phi_J$, $\epsilon = \epsilon_1 + \epsilon_2$, and

$$\mathcal{D}(x) = \frac{x^2}{(x^2 - \epsilon_2^2)} \frac{(x^2 - \epsilon^2)}{(x^2 - \epsilon_1^2)}, \quad (1.5)$$

$$Q(x) = \frac{1}{P(x)P(x+\epsilon)}, \quad P(x) = \prod_{\lambda=1}^N (x - a_\lambda). \quad (1.6)$$

The origin of this expression is reviewed in Appendix A. The parameters entering this integrals, namely $\epsilon_{1,2}$ and a_λ are taken to be real with a small positive imaginary part $i0$. The domain of integration above should be understood either as a real slice integration or equivalently, upon closing the integration in the upper-half plane, as a multiple contour integral. In Appendix A.4 we review how the residue evaluation of (1.4) reproduces the representation of the instanton partition function as sum over N -tuples of Young diagrams.

We emphasize that the precise form of $Q(x)$ and $\mathcal{D}(x)$ does not affect the derivation presented in this paper. This is the main reason why the generalization to quiver gauge theories is rather straightforward. On the other hand the presence of the factor $\frac{x^2}{x^2 - \epsilon_2^2}$ in $\mathcal{D}(x)$, which have a particularly singular limit for ϵ_2 small, will play a crucial role and will be responsible for the appearance of the dilogarithm function in the TBA.

TBA form

The claim of [18] is that the twisted superpotential, defined as

$$\mathcal{W} = \text{Limit}_{\epsilon_2 \rightarrow 0} (\epsilon_2 \log \mathcal{Z}), \quad (1.7)$$

can be written as the critical value of the following integral functional:

$$\mathcal{Y}[\rho, \varphi] = \frac{1}{2} \int \frac{dx}{2\pi i} \frac{dy}{2\pi i} \rho(x) G(x-y) \rho(y) + \int \frac{dx}{2\pi i} \left[\rho(x) \varphi(x) + \text{Li}_2(qQ(x)e^{-\varphi(x)}) \right], \quad (1.8)$$

which is

$$\mathcal{W} = \text{Crit}_{\rho, \varphi} [\mathcal{Y}(\rho, \varphi)] = \int \frac{dx}{2\pi i} \left[-\frac{1}{2} \varphi(x) \log(1 - qQ(x)e^{-\varphi(x)}) + \text{Li}_2(qQ(x)e^{-\varphi(x)}) \right], \quad (1.9)$$

where $\varphi(x)$ is the solution of the following TBA-like equation

$$\varphi(x) = \int \frac{dy}{2\pi i} G(x-y) \log(1 - qQ(y)e^{-\varphi(y)}). \quad (1.10)$$

$Q(x)$ is defined in (1.6), and the propagator $G(x)$ is related to $\mathcal{D}(x)$ as

$$G(x) = \text{Limit}_{\epsilon_2 \rightarrow 0} \frac{\mathcal{D}(x) - 1}{\epsilon_2} = \frac{d}{dx} \log \left(\frac{x + \epsilon_1}{x - \epsilon_1} \right). \quad (1.11)$$

2 Mayer-Cluster expansion

As mentioned in [18], the contour integral form of the instanton partition function (1.4) can be interpreted (for each k) as the partition function of a one dimensional non-ideal gas of particles ϕ_1, \dots, ϕ_k , subject to an external potential and a pair-wise interaction potential respectively given by

$$U^{\text{ext}}(x) = -\log \left(Q(x) \frac{\epsilon}{\epsilon_1 \epsilon_2} \right), \quad V^{\text{int}}(x) = -\log(\mathcal{D}(x)). \quad (2.1)$$

Upon summing over the number of particles k , the instanton partition function takes the form of a grand canonical partition function. The free energy of this gas can be studied by Mayer expansion techniques [48] (see for example [52] for a nice introduction), as pointed out in [18]. In this section, we will perform such kind of expansion in full details. The limit of $\epsilon_2 \rightarrow 0$ appears to be rather subtle. In order to perform this limit we need to face the problem of studying the leading behavior of some multiple integral where a parameter is small. It turns out that this can be conveniently studied by employing the method of expansion by regions [49] discussed in Appendix C. We introduce this method to provide a unified framework to study certain integrals, but we stress that all the result of this section are also obtained without employing such technique.

In order to analyze the behavior of the partition function (1.4) in the limit in which ϵ_2 is small, it is convenient to split the function $\mathcal{D}(x)$, see (1.5), into two parts:

$$\mathcal{D}(x) = \frac{x^2}{x^2 - \epsilon_2^2} \tilde{\mathcal{D}}(x), \quad \tilde{\mathcal{D}}(x) = \frac{x^2 - \epsilon^2}{x^2 - \epsilon_1^2}. \quad (2.2)$$

The reason of such decomposition is as follows. The factor $\frac{x^2}{x^2 - \epsilon_2^2}$ corresponds to a pair-wise interaction which is strong and attractive at distances of order ϵ_2 and rapidly decreases at large distances. The remaining factor $\tilde{\mathcal{D}}(x)$ corresponds to a pair-wise interaction which is different from zero only at distances of order ϵ_1 . Thus (2.2) corresponds to splitting the potential into short- and long-range parts. The natural candidate to study such kind of potentials is the so-called *iterated* Mayer expansion [50]. As we will see in Section 2.3, this effectively creates a new grand canonical partition function whose “new particles” correspond to clusters of the original particles. We will start by considering some simplified situation.

2.1 Only long range interactions

Let us first consider a simplified version of (1.4) without the factors² $\frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2}$, i.e

$$\mathcal{Z}_{\text{Long}} := \sum_{k=0}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \frac{Q(\phi_I)}{\epsilon_2} \prod_{1 \leq I < J \leq k} \tilde{\mathcal{D}}(\phi_{IJ}). \quad (2.3)$$

The basic idea of Mayer expansion, see e.g. [52], is to introduce the function \tilde{f}_{IJ} as

$$\tilde{\mathcal{D}}(\phi_{IJ}) = 1 + \tilde{f}_{IJ}, \quad (2.4)$$

and expand the interaction products as

$$\prod_{1 \leq I < J \leq k} (1 + \tilde{f}_{IJ}) = 1 + \sum_{I < J} \tilde{f}_{IJ} + \sum_{I < J, I' < J'} \tilde{f}_{IJ} \tilde{f}_{I'J'} + \dots \quad (2.5)$$

Each monomial in the right hand side of this equation can be visualized as a graph (not necessarily connected) on the set $[k] = \{1, 2, \dots, k\}$. More precisely, each particle in $\{1, 2, \dots, k\}$ corresponds to a vertex, and to each factor \tilde{f}_{IJ} we associate an edge between particle I and J in the corresponding graph. More explicitly,

$$\tilde{f}_{IJ} := \begin{array}{c} \text{---} \\ | \\ \bullet \quad \bullet \end{array} \quad (2.6)$$

As an example for the $k = 3$ case, we have the expansion in terms of graphs shown in Figure 1. It is clear from the left hand side of (2.5) that there are no multiple edges between two vertices, or edges connecting one vertex to itself.

²We also set $\frac{\epsilon}{\epsilon_1}$ to one.

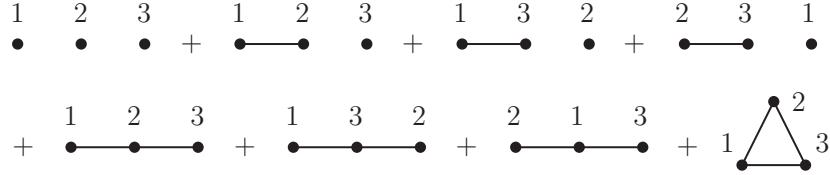


Figure 1: *The graph expansion for the $k = 3$ case. The second line contains all connected graphs.*

This expansion is particularly useful, as the logarithm of the grand canonical partition function can be formally given as a sum over *connected* graphs [48] (see for example Appendix A of [52] for a simple derivation):

$$\log \mathcal{Z}_{\text{Long}} = \sum_{k=1}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \frac{Q(\phi_I)}{\epsilon_2} \sum_{g \in \mathcal{G}_c^{[k]}} \prod_{e(I,J) \in g} \tilde{f}_{IJ}. \quad (2.7)$$

Here $\mathcal{G}_c^{[k]}$ denotes the collection of connected graphs on the set $[k]$ and $e(I, J)$ belongs to the graph g if there is an edge between the vertices I and J . Graphs up to four points are shown in Figure 2.

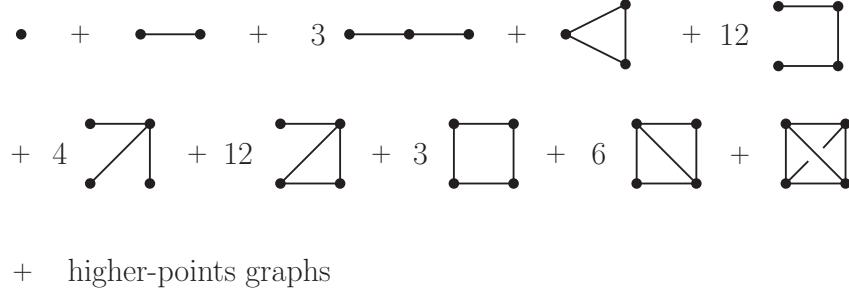


Figure 2: *Connected graphs in the Mayer expansion. The coefficients correspond to how many different graphs of the given topology belong to $\mathcal{G}_c^{[k]}$. They are obtained from one another by relabeling the vertices.*

We stress that (2.7) is an exact relation as a formal power series³ in q . In the limit of small ϵ_2 one has

$$\tilde{f}_{IJ} = \epsilon_2 G(\phi_{IJ}) + \mathcal{O}(\epsilon_2^2), \quad (2.8)$$

³The interesting question of convergence of the Mayer expansion can be addressed in various ways, see e.g. [52]. Here, we will not consider this problem or the analog convergence issue for the solution to the TBA.

where G is defined in (1.11). As each factor of \tilde{f} contributes one power of ϵ_2 , the leading contribution to the sum in (2.7) is given only by connected tree graphs $\mathcal{T}_c^{[k]}$. Such graphs have the minimal number of edges ($k - 1$) among the connected graphs. The sum of these tree graphs can be shown to be convergent. Collecting the powers of ϵ_2 we conclude that

$$\log \mathcal{Z}_{\text{Long}} = \frac{1}{\epsilon_2} \sum_{k=1}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} Q(\phi_I) \sum_{g \in \mathcal{T}_c^{[k]}} \prod_{e(I,J) \in g} \tilde{f}_{IJ} + \mathcal{O}(\epsilon_2^0). \quad (2.9)$$

Graphs up to five points are shown in Figure 3. Notice that the leading behaviour of the free energy (2.9) is at order $\frac{1}{\epsilon_2}$. This fact may be not obvious from the definition (2.3) where for each k the leading contribution is proportional to $\frac{1}{\epsilon_2^k}$. We will see that the similar behaviour applies to the more complicated situations analized in the following.

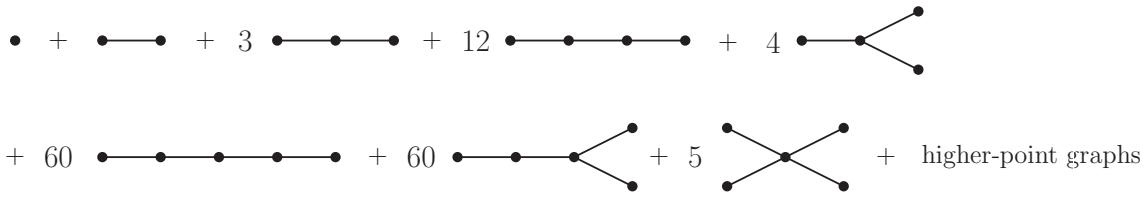


Figure 3: *Connected tree graphs up to five-point. The coefficients corresponds to how many different graphs of the given topology belong to $\mathcal{T}_c^{[k]}$, which are related to each other by relabeling the vertices.*

2.2 Only short range interactions

Next we study another simplified version of (1.4). Namely, we set $\tilde{\mathcal{D}}$, defined in (2.2), to zero and consider

$$\mathcal{Z}_{\text{Short}} := \sum_{k=0}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \frac{Q(\phi_I)}{\epsilon_2} \prod_{1 \leq I < J \leq k} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2}. \quad (2.10)$$

The main achievement of this subsection is to show that in the limit of small ϵ_2 the logarithm of (2.10) is given by

$$\log \mathcal{Z}_{\text{Short}} = \frac{1}{\epsilon_2} \int \frac{d\phi_0}{2\pi i} \text{Li}_2[qQ(\phi_0)] + \mathcal{O}(\epsilon_2^0). \quad (2.11)$$

We will see that, as opposed to the case of long range interactions considered in the previous subsection, the right hand side of (2.11) is not given by summing over tree graphs but by

a single local term. As will be explained in details in Section 2.3, this is the underlying mechanism by which the short-range interactions turn a group of particles into a single effective particle. The result (2.11) explains why the dilogarithm function appears in the TBA. It will also be an essential input in the all order proof presented in Section 3.2. Considering its importance, we will prove (2.11) in two different ways.

2.2.1 Dilogarithm from a sum over residues

In this subsection we will prove (2.11) by direct evaluating (2.10) and (2.11) as a sum over residues. We find it instructive to first consider the partition function for the $U(1)$ gauge theory, i.e. we evaluate (2.10) for

$$Q(x) = \frac{1}{(x - a - i0)(x - a + \epsilon + i0)}, \quad (2.12)$$

compare to (1.6). In (2.12) a and ϵ are real and we wrote explicitly the $i0$ prescription. For each k in (2.10), there is only one residue (up to permutation of the integration variables) in the upper half plane, compare to the general discussion in Appendix A.4. It is given by $(\phi_1, \dots, \phi_k) = (a, a + \epsilon_2, \dots, a + (k - 1)\epsilon_2)$. It follows that

$$\mathcal{Z}_{\text{Short}}^{U(1)} = \sum_{k=0}^{\infty} \frac{q^k}{k!} \frac{1}{\epsilon_2^k} \prod_{I=1}^k \frac{1}{\epsilon_1 + I\epsilon_2}. \quad (2.13)$$

One can also directly calculate the one dimensional integral

$$\int \frac{d\phi_0}{2\pi i} Q(\phi_0)^\ell = \frac{(-1)^{\ell+1}}{\epsilon^{2\ell-1}} \binom{2\ell-2}{\ell-1}, \quad (2.14)$$

where $Q(x)$ is given in (2.12). Using the identity (B.6) one concludes the the logarithm of (2.13) in the $\epsilon_2 \rightarrow 0$ limit is indeed given by (2.11) with (2.14).

For the more general $U(N)$ theory, it is easy to classify the poles contributing to (2.10) following the same reasoning as in Appendix A.4. For fixed k , residues are classified, up to permutation of the k particles, by a set of integers $\{s_1, s_2, \dots, s_N\}$ such that $\sum_\lambda s_\lambda = k$. The corresponding pole is given by

$$\phi_{\kappa_\lambda + I_\lambda} = a_\lambda + (I_\lambda - 1)\epsilon_2, \quad I_\lambda = 1, \dots, s_\lambda, \quad \kappa_\lambda = \sum_{\mu=1}^{\lambda-1} s_\mu, \quad (2.15)$$

and $\lambda = 1, \dots, N$. As opposed to the full partition function (1.4), for which residues are classified by N -tuple of Young tableaux with a total number of boxes equal to k , in the simplified integrals (2.10) only Young tableaux with one row, whose length is denoted by

s_λ , contribute. It is straightforward to calculate and collect all residues, see Appendix B, to obtain

$$\mathcal{Z}_{\text{Short}} = \sum_{s_1, \dots, s_N \geq 1} \prod_{\lambda=1}^N \left(\frac{q^{s_\lambda} \mathfrak{R}_{s_\lambda}^{(\lambda)}}{s_\lambda! \epsilon_2^{s_\lambda}} \right) \prod_{\lambda < \lambda'} \prod_{I_\lambda, I_{\lambda'}} \frac{(a_{\lambda\lambda'} + (I_\lambda - I_{\lambda'})\epsilon_2)^2}{(a_{\lambda\lambda'} + (I_\lambda - I_{\lambda'})\epsilon_2)^2 - \epsilon_2^2}, \quad (2.16)$$

where

$$\mathfrak{R}_s^{(\lambda)} := \prod_{I=1}^s \frac{Q^{(\lambda)}(a_{\lambda_1} + (I-1)\epsilon_2)}{\epsilon_1 + I\epsilon_2}, \quad Q^{(\lambda)}(x) = \frac{1}{P^{(\lambda)}(x)P^{(\lambda)}(x+\epsilon)}, \quad (2.17)$$

and $P^{(\lambda)}(x) = \prod_{\mu \neq \lambda} (x - a_\mu)$. The structure of the result (2.16)-(2.17) represents a simple generalization of the $U(1)$ computation (2.13). The $\epsilon_2 \rightarrow 0$ limit of the logarithm of (2.16) can be readily obtained using the relation (B.7). One recognizes that

$$\log \mathcal{Z}_{\text{Short}} = \frac{1}{\epsilon_2} \sum_{\lambda=1}^N \sum_{\ell=1}^{\infty} \frac{q^\ell}{\ell^2} \text{Res}_{\phi_0=a_\lambda} [Q^\ell(\phi_0)] + \mathcal{O}(\epsilon_2^0). \quad (2.18)$$

This result coincides with the evaluation of (2.11) by residues. This calculation provides a direct proof of (2.11).

2.2.2 By Mayer expansion and separation of regions

We now calculate (2.10) for small ϵ_2 by first applying Mayer expansion and then the so-called method of expansion by regions. The first step is to decompose the interaction factor as

$$\frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} = 1 + f_{IJ}, \quad f_{IJ} := \frac{\epsilon_2^2}{\phi_{IJ}^2 - \epsilon_2^2}. \quad (2.19)$$

In a similar way as (2.7), the logarithm of (2.10) is then given by

$$\log \mathcal{Z}_{\text{Short}} = \sum_{k=1}^{\infty} \frac{q^k}{k! \epsilon_2^k} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} Q(\phi_I) \sum_{g \in \mathcal{G}_c^{[k]}} \prod_{e(I,J) \in g} f_{IJ}. \quad (2.20)$$

To show that this gives (2.11), we need to evaluate the integrals entering (2.20) as a Laurant series in ϵ_2 . We are actually interested only in the terms with leading negative powers of ϵ_2 . We use a powerful method, usually applied to the evaluation of Feynman integrals, called *expansion by regions* introduced in [49]. The method goes as follows: (1) divide the integration domain into regions and expand the integrand in a Taylor series in small parameters, (2) extend the integration to the full domain of integration, (3) set to zero scaleless integrals. This method turns out to be particularly efficient to calculate the leading term for the integrals (2.20). Concerning these contributions, we do not need to apply the

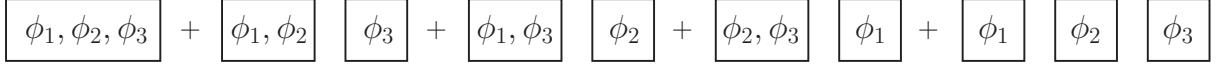


Figure 4: *The decomposition of the contribution for $k = 3$ case. Each square represents a cluster labeled by corresponding particles in the cluster.*

somewhat subtle point (3) above. A more detailed discussion of this method will be given in Appendix C.

To identify the set of relevant regions, we note that the Taylor expansion of the interaction term $\frac{\epsilon_2^2}{\phi_{IJ}^2 - \epsilon_2^2}$ starts at order ϵ_2^2 except for the *region* in which ϕ_{IJ} is of order ϵ_2 . It is then natural to expect that regions are classified as follows. Let $\mathcal{B}^{[k]}$ denotes the set of grouping of k labeled particles in clusters, see figure 4 for the $k = 3$ case. For each grouping we define the corresponding region as

$$\begin{cases} |\phi_{IJ}| \sim \epsilon_2 & \text{if } I, J \text{ are in the same cluster,} \\ |\phi_{IJ}| \gg \epsilon_2 & \text{if } I, J \text{ are in different clusters.} \end{cases} \quad (2.21)$$

The next step is to Taylor expand the integrand in each region. We denote by $\mathbb{T}^{(b)} [\dots]$ the operation of Taylor expaning \dots in the region corresponding to the grouping $b \in \mathcal{B}^{[k]}$. We have

$$\mathbb{T}^{(b)} [f_{IJ}] = \begin{cases} \mathcal{O}(1) & \text{if } I, J \text{ are in the same cluster,} \\ \mathcal{O}(\epsilon_2^2) & \text{if } I, J \text{ are in different clusters,} \end{cases} \quad (2.22)$$

and

$$\mathbb{T}^{(b)} \left[\prod_{I=1}^k Q(\phi_I) \right] = \prod_{\mathbf{Y} \in b} Q(\phi_{\mathbf{Y}})^{|\mathbf{Y}|} + \dots \quad \phi_{\mathbf{Y}} := \frac{1}{|\mathbf{Y}|} \sum_{I \in \mathbf{Y}} \phi_I, \quad (2.23)$$

the product on the right hand side runs over the clusters in the grouping b and $|\mathbf{Y}|$ denotes the number of particles in the cluster \mathbf{Y} . From (2.22)-(2.23) it is clear that, for each k , the leading contribution to (2.20) comes from the region in which all ϕ_I are in the same cluster. Indeed, if there were two or more clusters, the f_{IJ} factor, with I, J in different clusters, would decrease the power of ϵ_2 , compare to (2.22). An explicit example of this expansion of $k = 3$ is given in appendix C. As we will shortly see, the leading contribution to (2.20) starts at order $1/\epsilon_2$ for each k . This fact is not obvious from the form (2.20).

By further separating the integration over the “center of cluster” coordinate \bar{x} as

$$\int \prod_{I=1}^s dx_I F(x_1, \dots, x_s) = \int d\bar{x} \int \prod_{I=1}^s dx'_I \delta \left(\frac{1}{k} \sum_{I=1}^s x'_I \right) F(\bar{x} + x'_1, \dots, \bar{x} + x'_s), \quad (2.24)$$

where we applied the change of variables

$$\bar{x} := \frac{1}{k} \sum_{I=1}^s x_s, \quad x'_I := x_I - \bar{x}, \quad (2.25)$$

one can rewrite (2.20) as

$$\log \mathcal{Z}_{\text{Short}} = \frac{1}{\epsilon_2} \sum_{k=1}^{\infty} \frac{q^k}{k!} \int d\phi_0 Q^k(\phi_0) \mathcal{J}_k + \mathcal{O}(\epsilon_2^0) \quad (2.26)$$

where

$$\mathcal{J}_k := \frac{1}{\epsilon_2^{k-1}} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \delta \left(\frac{1}{k} \sum_{I=1}^k \phi_I \right) \sum_{g \in \mathcal{G}_c^{[k]}} \prod_{e(I,J) \in g} f_{IJ}. \quad (2.27)$$

Notice that in (2.27) all connected graphs, independently on the number of edges, contribute to the leading term. The next observation to be made is that

$$\text{if } g \text{ is not connected} \quad \frac{1}{\epsilon_2^{k-1}} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \delta \left(\frac{1}{k} \sum_{I=1}^k \phi_I \right) \prod_{e(I,J) \in g} f_{IJ} = 0. \quad (2.28)$$

The validity of this statement can be easily argued as follows. For each connected component in g we can define its center as the average of the $\{\phi_I\}$ in that connected component. The integrand in (2.28) does not depend on the distance between the centers. As the integral is calculated by residue, it trivially vanishes in this case as, after integrating trivially the delta function, the $(k-1)$ -dimensional residue is absent. The identity (2.28) implies that in (2.27), we can replace the sum over connected graphs with the sum over all graphs. Finally using the relation

$$\sum_{g \in \mathcal{G}^{[k]}} \prod_{e(I,J) \in g} f_{IJ} = \prod_{1 \leq I < J \leq k} (1 + f_{IJ}), \quad (2.29)$$

where $\mathcal{G}^{[k]}$ is the set of all graphs on $[k]$, we conclude that $\mathcal{J}_k = \mathcal{I}_k$, where, using (2.19),

$$\mathcal{I}_k := \frac{1}{\epsilon_2^{k-1}} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \delta \left(\frac{1}{k} \sum_{I=1}^k \phi_I \right) \prod_{1 \leq I < J \leq k} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} = \frac{1}{2\pi i} \frac{k!}{k^2}. \quad (2.30)$$

Notice that the evaluation of this integral is exact, see [10] for a derivation. Using this result we recognize that (2.26) is equal to

$$\log \mathcal{Z}_{\text{Short}} = \frac{1}{\epsilon_2} \sum_{k=1}^{\infty} \frac{q^k}{k^2} \int \frac{d\phi_0}{2\pi i} Q^k(\phi_0) + \mathcal{O}(\epsilon_2^0) = \frac{1}{\epsilon_2} \int \frac{d\phi_0}{2\pi i} \text{Li}_2[qQ(\phi_0)] + \mathcal{O}(\epsilon_2^0). \quad (2.31)$$

2.3 Expansion of the instanton partition function

We are now ready to consider the full instanton partition function (1.4). We will combine the considerations from the previous sections. In Section 2.1 we learned that the $\epsilon_2 \rightarrow 0$ limit singles out certain tree level graphs. In Section 2.2 we learned that the factors $\frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2}$ produce the effect of combining particles into clusters. To exploit the combination of these two mechanisms in the most transparent way we find it convenient to use the so-called iterated Mayer expansion advocated at the beginning of our analysis. We start by reviewing this expansion. Based on this iterative expansion, by combining the discussion from the two previous sections, one obtains a tree graph expansion for the full instanton partition function as given in Section 2.3.2. In Appendix D we show that the same result can be obtained by first applying the method of expansion by regions to the original partition function and then exploiting some combinatorics to conclude that only certain connected graphs contribute to the free energy.

2.3.1 Iterated Mayer expansion

We start with a review of the iterated Mayer expansion [50]. Consider the partition function

$$\mathcal{Z} = \sum_{k=0}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k d\phi_I A(\phi_I) \prod_{1 \leq I < J \leq k} (1 + \mathcal{F}_{IJ}) , \quad 1 + \mathcal{F}_{IJ} = e^{-(V_S^{\text{int}}(\phi_{IJ}) + V_L^{\text{int}}(\phi_{IJ}))} , \quad (2.32)$$

where, as in (2.2), we split the pair-wise interaction potential in a short and long range part. Next we introduce

$$\mathfrak{g}_{\mathbf{Y}} = \prod_{\substack{I, J \in \mathbf{Y} \\ I < J}} \left(1 + e^{-V_L^{\text{int}}(\phi_{IJ})} \right) , \quad 1 + f_{IJ} = e^{-V_S^{\text{int}}(\phi_{IJ})} , \quad (2.33)$$

$$1 + \mathfrak{f}_{\mathbf{Y}_a, \mathbf{Y}_b} = \prod_{I_a \in \mathbf{Y}_a} \prod_{I_b \in \mathbf{Y}_b} e^{-V_L^{\text{int}}(\phi_{I_a I_b})} , \quad a \neq b \quad (2.34)$$

where \mathbf{Y} denotes a set of particles or, in other words, a cluster. Iterated Mayer expansion is the statement that the free energy can be expanded as

$$\log \mathcal{Z} = \sum_{k=1}^{\infty} \frac{q^k}{k!} \sum_{\ell \geq 0} \sum_{\{\mathbf{Y}_1, \dots, \mathbf{Y}_\ell\} \in \mathcal{B}_\ell^{[k]}} \int \prod_{I=1}^k d\phi_I \mathfrak{S}_{\{\mathbf{Y}_1, \dots, \mathbf{Y}_\ell\}} , \quad (2.35)$$

where

$$\mathfrak{S}_{\{\mathbf{Y}_1, \dots, \mathbf{Y}_\ell\}} = \left(\prod_{a=1}^{\ell} \mathfrak{g}_{\mathbf{Y}_a} S_{\mathbf{Y}_a} \right) \sum_{g \in \mathcal{G}_c^{[\ell]}} \prod_{e(a, b) \in g} \mathfrak{f}_{\mathbf{Y}_a, \mathbf{Y}_b} , \quad (2.36)$$

$$S_{\mathbf{Y}} = \prod_{I=1}^k A(\phi_I) \sum_{g \in \mathcal{G}_c^{\mathbf{Y}}} \prod_{e(I,J) \in g} f_{IJ}, \quad (2.37)$$

The sum in (2.35) is taken over $\mathcal{B}_{\ell}^{[k]}$: groupings of k (labeled) particles into ℓ clusters. As the notation may need some time to be digested, in Appendix B.3 we spell out the definitions for $k = 2, 3$. Note that although the cluster here is in a different context, the picture of grouping is similar to that used before in expansion by regions, see for example figure 4 for $k = 3$ case.

2.3.2 Iterated Mayer expansion for the Nekrasov partition function

We can apply the iterated Mayer expansion reviewed in Section 2.3.1 to the full partition function, compare (2.32) to (1.4), (2.2). Once this is done we need to evaluate the leading contribution for small ϵ_2 to integrals of the type (2.35). The crucial observation is that the Taylor expansion $\tilde{D}(\phi_{IJ}) - 1 \sim \epsilon_2 G(\phi_{IJ})$ is valid everywhere in the domain of integration⁴. This immediately implies that in the limit of small ϵ_2 one can write, see definition (2.36),

$$\mathfrak{S}_{\{\mathbf{Y}_1, \dots, \mathbf{Y}_{\ell}\}} = \epsilon_2^{\ell-1} \prod_{a=1}^{\ell} S_{\mathbf{Y}_a} \sum_{g \in \mathcal{T}_c^{[\ell]}} \prod_{e(a,b) \in g} \prod_{\substack{I_a \in \mathbf{Y}_a \\ I_b \in \mathbf{Y}_b}} G(\phi_{I_a I_b}) + \dots \quad (2.38)$$

where \dots refers to next to leading contributions in ϵ_2 . We point out that, in analogy with (2.9), only connected *tree graphs* contributes to (2.38). As opposed to (2.9), now they are tree graphs on the set $[\ell]$ of ℓ clusters rather than the set of k fundamental particles. Notice that we did not expand the $S_{\mathbf{Y}}$ factor. The integration in (2.35) is still over k variables.

The next step is to explicitly perform the integration over the distances of particles within the same cluster. This turns out to be essentially the same as in Section 2.2.2. For each cluster \mathbf{Y}_a we introduce a ‘‘center of cluster’’ coordinate $\bar{\phi}_a$ as in (2.24). It is clear that the f_{IJ} entering the factors $S_{\mathbf{Y}_a}$, see (2.37), are independent of the ‘‘center of cluster’’ coordinates $\bar{\phi}_a$. Using this observation and (2.38) we write

$$\begin{aligned} & \int \prod_{I=1}^k d\phi_I \mathfrak{S}_{\{\mathbf{Y}_1, \dots, \mathbf{Y}_{\ell}\}} \\ &= \frac{1}{\epsilon_2} \int \prod_{a=1}^{\ell} \left[d\bar{\phi}_a d\mu_{\mathbf{Y}_a} \prod_{I_a \in \mathbf{Y}_a} Q(\bar{\phi}_a + \phi_{I_a}) \right] \sum_{g \in \mathcal{T}_c^{[\ell]}} \prod_{e(a,b) \in g} \prod_{\substack{I_a \in \mathbf{Y}_a \\ I_b \in \mathbf{Y}_b}} G(\bar{\phi}_{ab} + \phi_{I_a I_b}) + \dots \end{aligned} \quad (2.39)$$

where

$$d\mu_{\mathbf{Y}} := \frac{1}{\epsilon_2^{|\mathbf{Y}|-1}} \prod_{I \in \mathbf{Y}_a} \frac{d\phi_I}{2\pi i} \delta \left(\frac{1}{|\mathbf{Y}|} \sum_{I \in \mathbf{Y}} \phi_I \right) \sum_{g \in \mathcal{G}_c^{\mathbf{Y}}} \prod_{e(I,J) \in g} f_{IJ}. \quad (2.40)$$

⁴ In the language of the method of expansion by regions we would say that the leading term in the Taylor expansion of $\tilde{D}(\phi_{IJ}) - 1$ is the same in *any region*.

As extensively discussed in the previous sections, the leading contribution in ϵ_2 of integrals of the type (2.39), can be obtained by neglecting the deviation from the center of cluster coordinate in the functions $Q(x)$ and $G(x)$. Doing so, (2.39) becomes

$$\int \prod_{I=1}^k d\phi_I \mathfrak{S}_{\{\mathbf{Y}_1, \dots, \mathbf{Y}_\ell\}} = \frac{1}{\epsilon_2} \int \prod_{a=1}^\ell d\bar{\phi}_a Q^{|\mathbf{Y}_a|}(\bar{\phi}_a) \mathcal{J}_{|\mathbf{Y}_a|} \sum_{g \in \mathcal{T}_c^{[\ell]}} \prod_{e(a,b) \in g} |\mathbf{Y}_a| |\mathbf{Y}_b| G(\bar{\phi}_{ab}) + \dots \quad (2.41)$$

\mathcal{J}_k was defined in (2.27) and computed in (2.30). For convenience we recall it here

$$\mathcal{J}_{|\mathbf{Y}_a|} = \int d\mu_{\mathbf{Y}_a} = \frac{1}{2\pi i} \frac{|\mathbf{Y}_a|!}{|\mathbf{Y}_a|^2}. \quad (2.42)$$

In the limit of small ϵ_2 , the summands in (2.35) depend only on the sizes $n_1, \dots, n_\ell \geq 1$ of the clusters corresponding to $\mathbf{b} \in \mathcal{B}^{[k]}$. Converting the sum in (2.35) from a sum over groupings of k particles to a sum over the number of clusters ℓ and their sizes n_1, \dots, n_ℓ produces a factor $\frac{k!}{\ell! n_1! \dots n_\ell!}$. Assembling the pieces together we finally arrive at the main result

$$\lim_{\epsilon_2 \rightarrow 0} \epsilon_2 \log \mathcal{Z} = \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \sum_{n_1, \dots, n_\ell \geq 1} \int \prod_{a=1}^\ell \frac{d\phi_a}{2\pi i} \frac{q^{n_a} Q^{n_a}(\phi_a)}{n_a^2} \sum_{g \in \mathcal{T}_c^{[\ell]}} \prod_{e(a,b) \in g} n_a n_b G(\phi_{ab}), \quad (2.43)$$

where we replaced $\bar{\phi}_a$ with ϕ_a . We emphasize once again that the sum has been rearranged from a sum over particles, weighted by q , to a sum over clusters. More precisely, for each ℓ there are ℓ clusters and $\ell - 1$ G -factors. For $\ell = 1$ one immediately recovers the dilogarithm

$$\int \frac{d\phi}{2\pi i} \text{Li}_2(q Q(\phi)). \quad (2.44)$$

For $\ell = 2$ one has

$$\frac{1}{2} \int \frac{d\phi_1}{2\pi i} \frac{d\phi_2}{2\pi i} \mathcal{Q}_{\text{eff}}(\phi_1) \mathcal{Q}_{\text{eff}}(\phi_2) G(\phi_1 - \phi_2), \quad \mathcal{Q}_{\text{eff}}(\phi) = \log(1 - q Q(\phi)). \quad (2.45)$$

The result (2.43) can be visualized diagrammatically in a simple way. For each ℓ in the sum one draws all three graphs on the set $\{1, 2, \dots, \ell\}$. To each node a of the graph is associated an integer $n_a \geq 1$. Once the diagrams are drawn, the corresponding integrals can be written using the following ‘‘Feynann rules’’

- Vertex

$$\boxed{n_a} = \frac{q^{n_a}}{n_a^2} \int \frac{d\phi_a}{2\pi i} Q^{n_a}(\phi_a) \quad (2.46)$$

- Propagator

$$\begin{array}{c} \boxed{n_a} \\ \hline \boxed{n_b} \end{array} = n_a n_b G(\phi_{ab}). \quad (2.47)$$

Here we used a dashed square to indicate that the propagator is associated to the edge only.

The graphs contributing to $\ell = 1, 2, 3$ are given by

$$\ell = 1 : \sum_{n_1=1}^{\infty} \boxed{n_1} \quad (2.48)$$

$$\ell = 2 : \sum_{n_1, n_2=1}^{\infty} \boxed{n_1} - \boxed{n_2} \quad (2.49)$$

$$\ell = 3 : \sum_{n_1, n_2, n_3=1}^{\infty} \boxed{n_1} - \boxed{n_2} - \boxed{n_3} + \boxed{n_1} - \boxed{n_3} - \boxed{n_2} + \boxed{n_2} - \boxed{n_1} - \boxed{n_3} \quad (2.50)$$

The first two graphs correspond to (2.44) and (2.45) respectively. Notice that each integral obtained by applying the Feynman rules, should be multiplied by an overall $\frac{1}{\ell}$ factor, which is explicit in (2.43). Graphs with more vertices, upon dressing the vertices with positive integers, are as in Figure 3.

3 Derivation of the TBA

In this section we show that the logarithm of the instanton partition function in the $\epsilon_2 \rightarrow 0$ limit, whose structure have been studied in last section, satisfies TBA equations. We will first provide a perturbative check to a very high order in the instanton counting parameter q , and then present an all order proof based on rewriting the instanton partition function as a $(0+1)$ -dimensional path integral.

3.1 Perturbative expansion

Let us recall the expression of the superpotential coming from the TBA

$$\mathcal{W} = \int \frac{dx}{2\pi i} \left[-\frac{1}{2} \varphi(x) \log(1 - qQ(x)e^{-\varphi(x)}) + \text{Li}_2(qQ(x)e^{-\varphi(x)}) \right], \quad (3.1)$$

where

$$\varphi(x) = \int \frac{dy}{2\pi i} G(x-y) \log(1 - qQ(y)e^{-\varphi(y)}), \quad (3.2)$$

compare to (1.9) and (1.10). The basic observation is that TBA equation has a natural expansion in terms of tree graphs, therefore one can compare it directly with the Mayer expansion of the instanton partition function (2.43).

One writes φ and \mathcal{W} as

$$\varphi(x) = \sum_{k=1}^{\infty} q^k \varphi_k(x), \quad \mathcal{W} = \sum_{k=1}^{\infty} q^k \mathcal{W}_k. \quad (3.3)$$

φ can be solved recursively via the TBA equation, for example up to the first two orders

$$\varphi_1(x) = - \int \frac{dy}{2\pi i} G(x-y) Q(y), \quad (3.4)$$

$$\varphi_2(x) = -\frac{1}{2} \int \frac{dy}{2\pi i} G(x-y) Q(y)^2 - \int \frac{dy}{2\pi i} \frac{dz}{2\pi i} G(x-y) G(y-z) Q(y) Q(z). \quad (3.5)$$

Substituting the expression for φ in \mathcal{W} and collecting the terms at a given order of q , one obtains, here up to order q^3

$$\mathcal{W}_1 = \int \frac{dx}{2\pi i} Q(x), \quad (3.6)$$

$$\mathcal{W}_2 = \frac{1}{4} \int \frac{dx}{2\pi i} Q(x)^2 + \frac{1}{2} \int \frac{dx}{2\pi i} \frac{dy}{2\pi i} Q(x) G(x-y) Q(y), \quad (3.7)$$

$$\begin{aligned} \mathcal{W}_3 = & \frac{1}{9} \int \frac{dx}{2\pi i} Q(x)^3 + \frac{1}{2} \int \frac{dx}{2\pi i} \frac{dy}{2\pi i} Q(x)^2 G(x-y) Q(y) \\ & + \frac{1}{2} \int \frac{dx dy dz}{(2\pi i)^3} Q(x) G(x-y) Q(y) G(y-z) Q(z). \end{aligned} \quad (3.8)$$

One can see that the expansion of \mathcal{W} has a natural interpretation in terms of connected tree diagrams, with $G(x-y)$ as propagators and $Q(x)^i$ as vertices. This is the same structure as the one obtained starting from the integral representation of the instanton partition function, see (2.43). The non-trivial thing to show is that not only the structure is similar but that the coefficients actually match.

By staring at their form, it is not obvious that the Mayer expansion (2.43) and the TBA actually produce the same result. In particular to obtain the contribution of a single graph in (2.43) one needs to take into account a large number of terms coming from the φ expansion. This procedure can be straightforward implemented on a computer. We have checked up to q^7 order and found perfect agreement with the Mayer-Cluster expansion of the instanton partition function. It is remarkable that these two different structures are actually equivalent with each other. In the next subsection, we will see the equivalence of the two expansions can be proven to all order in q .

3.2 TBA as saddle-point of a $(0+1)$ -dimensional path integral

As we will shortly review, there is a nice way to rewrite a grand canonical partition function of the form (1.4) as a path integral, based on a trick of Polyakov [51]. Path integrals are particularly well suited to be studied in the classical limit $\hbar \rightarrow 0$, identified with ϵ_2 in this section, where the path integral is obtained by evaluating the classical action on some solution to the equations of motion. In the case of the instanton partition function (1.4), this limit is more subtle as the two particle potential in the original statistical mechanical problem is singular in this limit. Using the considerations from section 2.2 we will be able to

disentangle such contributions and derive the TBA in a rather transparent way as a saddle point equation of an appropriate action.

Let us start by reviewing the steps presented in the appendix of [51]. Consider the partition function

$$Z = \sum_{k=0}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \prod_{I < J} e^{-V(\phi_{IJ})}. \quad (3.9)$$

The basic idea is to rewrite the interactions as

$$\prod_{I < J} e^{-V(\phi_{IJ})} = e^{\Delta(\frac{\delta}{\delta\varphi})} \left. \prod_{I=1}^k e^{-\varphi(\phi_I)} \right|_{\varphi \rightarrow 0}, \quad (3.10)$$

where

$$\Delta\left(\frac{\delta}{\delta\varphi}\right) = -\frac{1}{2} \int dx dy V(x-y) \frac{\delta}{\delta\varphi(x)} \frac{\delta}{\delta\varphi(y)}. \quad (3.11)$$

The operator e^Δ can be pulled out from the sum over the number of particles k to rewrite

$$Z = e^{\Delta(\frac{\delta}{\delta\varphi})} \left(\sum_{k=0}^{\infty} \frac{1}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} e^{-\varphi(\phi_I)} \right) \Big|_{\varphi=0} = e^{\Delta(\frac{\delta}{\delta\varphi})} \exp\left(\int \frac{d\phi_0}{2\pi i} e^{-\varphi(\phi_0)}\right) \Big|_{\varphi \rightarrow 0}. \quad (3.12)$$

All the information about the interaction in the original gas (3.9) is now encoded in the operator Δ . To translate the construction above into path integral language one observes the following equality

$$e^{\Delta(\frac{\delta}{\delta\varphi})} F[\varphi] \Big|_{\varphi=0} = \int [D\varphi D\rho] e^{\frac{1}{\hbar} [\hbar \Delta(\frac{i\rho}{2\pi\hbar}) + \int \frac{d\varphi}{2\pi i} \rho(x) \varphi(x)]} F[\varphi]. \quad (3.13)$$

This can be formally proved by Fourier transformation⁵. The measure in $[D\varphi D\rho]$ is defined in such a way that if $F[\varphi] = 1$ then the resulting path integral should also evaluate to 1. Notice that the relation (3.13) holds also when the operator Δ is not quadratic. If Δ is linear $\Delta = \int_x \gamma(x) \frac{\delta}{\delta\varphi(x)}$, the ρ integration produces a delta function and the resulting path integral gives $F[\gamma]$ as it should. Indeed in this case e^Δ is just the shift operator.

From the discussion in Section 2, it is rather clear that, in the study of the instanton partition function, we should apply the procedure just described in a slightly modified way. In particular we will factor out only the long range part of the potential (2.2), corresponding to $\tilde{\mathcal{D}}$, as we will now describe. The first step is to introduce

$$\Delta_L\left(\frac{\delta}{\delta\varphi}\right) = -\frac{1}{2} \int dx dy V_L(x-y) \frac{\delta}{\delta\varphi(x)} \frac{\delta}{\delta\varphi(y)}, \quad e^{-V_L(x)} = \tilde{\mathcal{D}}(x). \quad (3.15)$$

⁵The relation (3.13) is a path integral analogue of

$$e^{\Delta(\frac{\partial}{\partial x})} \mathbf{f}(x) \Big|_{x \rightarrow 0} = \int \frac{dp}{2\pi} e^{ipx + \Delta(ip)} \hat{\mathbf{f}}(p) \Big|_{x \rightarrow 0} = \int \frac{dp dy}{2\pi} e^{-ipy + \Delta(ip)} \mathbf{f}(y), \quad (3.14)$$

where $\hat{\mathbf{f}}(p)$ is introduced via Fourier transformation.

As above, one uses this operator to rewrite the long range interactions as in (3.10), and factor out their contribution from the sum over the instanton number k as follows

$$\mathcal{Z} = e^{\tilde{\Delta}(\frac{\delta}{\delta\varphi})} \mathbf{F}[\varphi] \Big|_{\varphi=0}, \quad \mathbf{F}[\varphi] = \sum_{k=0}^{\infty} \frac{1}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \frac{qQ(\phi_I) e^{-\varphi(\phi_I)}}{\epsilon_2} \prod_{I < J}^k \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2}. \quad (3.16)$$

Notice that we did not expand the short range interaction $\frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2}$, which still appears in $\mathbf{F}[\varphi]$. The reader may recognize that we already calculated this integral in the $\epsilon_2 \rightarrow 0$ limit in Section 2.2. More precisely $\mathbf{F}[\varphi]$ coincides with $\mathcal{Z}_{\text{Short}}$ defined in (2.10), upon replacing $Q(x)$ with $Q(x)e^{-\varphi(x)}$. Using the result (2.11) we can thus rewrite

$$\mathbf{F}[\varphi] = \exp \left(\frac{1}{\epsilon_2} \int \frac{d\phi_0}{2\pi i} \text{Li}_2[qQ(\phi_0) e^{-\varphi(\phi_0)}] + O(\epsilon_2^0) \right). \quad (3.17)$$

The appearance of the dilogarithm is quite remarkable. The derivation presented makes its origin from short range interactions completely transparent. The expression (3.16) with (3.17), is well suited for the semiclassical analysis in the $\epsilon_2 \rightarrow 0$ limit. Using (3.13) in this case, and recalling that $V_L(x) = -\epsilon_2 G(x) + \mathcal{O}(\epsilon_2^2)$ one immediately obtains

$$\mathcal{Z} = \int [D\varphi D\rho] \exp \left[\frac{1}{\epsilon_2} (\mathcal{Y}[\rho, \varphi] + O(\epsilon_2)) \right], \quad (3.18)$$

where $\mathcal{Y}[\rho, \varphi]$ is defined in (1.8) and we identified \hbar with ϵ_2 . In the limit of small ϵ_2 this path integral is calculated by the saddle point method⁶. The superpotential is thus obtained from the critical values of the functional $\mathcal{Y}[\rho, \varphi]$. As explained in Section 1.1, the TBA immediately follows. This completes the derivation.

3.2.1 Multi-particle interactions

Let us mention that the above discussion may be generalized to the cases containing “multi-particle” interactions, such as

$$\mathcal{Z} = \sum_{k=0}^{\infty} \frac{q^k}{k!} \int \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \frac{qQ(\phi_I)}{\epsilon_2} \prod_{I < J}^k \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} \prod_{m=2}^{\infty} \prod_{I_1 < \dots < I_m}^k e^{-V_m(\phi_{I_1}, \phi_{I_2}, \dots, \phi_{I_m})}. \quad (3.19)$$

One can similarly introduce

$$\Delta_m \left(\frac{\delta}{\delta\varphi} \right) = -\frac{1}{m!} \int d^m x V_m(x_1, \dots, x_m) \prod_{I=1}^m \frac{\delta}{\delta\varphi(x_I)}. \quad (3.20)$$

⁶Alternatively, one can also integrate out ρ exactly, as it appears quadratically in the action, before taking the $\epsilon_2 \rightarrow 0$ limit. The resulting path integral, in φ in the limit of small ϵ_2 , is dominant by the saddle point of φ , where the saddle point equation is exactly the TBA equation (1.10).

In analogy with (3.13), this can be brought to the following form

$$\mathcal{Z} = \int [D\rho D\varphi] \exp \left[- \sum_{m=2}^{\infty} \frac{(-1)^m}{m!} \int \frac{d^m x}{(2\pi i)^m} V_m(x_1, \dots, x_m) \prod_{I=1}^m \frac{\rho(x_I)}{\epsilon_2} + \int \frac{dx}{2\pi i \epsilon_2} \rho(x) \varphi(x) \right] \mathbf{F}[\varphi]. \quad (3.21)$$

Notice that if $m > 2$ one can no longer integrate out ρ exactly. The tri-particle interaction is directly related to certain gauge theories. The main example is given by quiver theories containing matter in the trifundamental representation of $U(2)^3$ [53]. The contour integral representation of the instanton partition function for this case can be found in [54]. Notice that one has $V_2 = -\epsilon_2 G + \mathcal{O}(\epsilon_2^2)$, $V_3 = \epsilon_2^2 G_3 + \mathcal{O}(\epsilon_2^3)$. This is the right scaling to produce a $\frac{1}{\epsilon_2}$ factor in the exponent of (3.21) which is crucial for the semiclassical analysis.

4 Generalization to quiver gauge theories

Until now the object of study has been the instanton partition function for pure $\mathcal{N} = 2$ Super-Yang-Mills with gauge group $SU(N)$, see (1.4). This partition function admits a natural generalization corresponding to other $\mathcal{N} = 2$ gauge theories. In order to specify an $\mathcal{N} = 2$ supersymmetric gauge theory the first step is to fix the following data: a gauge group G and a set of irreducible representations R of G corresponding to its matter content. Such datum gives rise to a consistent, or UV-complete, theory only if the beta function β_v associated to each simple factor of $G = \prod_v G_v$ satisfies $\beta_v \leq 0$. This requirement puts strong constraints on R and G . A classification of consistent theories can be found in the recent work [55]. In particular, with the exception of $G = SU(2)^V$ with R given by trifundamentals introduced in [53], each irreducible component of R is charged under at most two G_v . In the following we restrict to the case

$$G = \prod_{v=1}^V SU(N_v), \quad R = \sum_{v,w} c_{vw} (\square_v, \overline{\square}_w) + \sum_v n_v \square_v. \quad (4.1)$$

Theories in this family with vanishing beta functions falls into three classes, referred to as type I, II, II* in [56]. The matrix c in (4.1) is then identified with the adjacency matrix of the Dynkin diagram of type ADE (type I) and $\widehat{A}\widehat{D}\widehat{E}$ (type II and II*). Any other consistent theory in the family (4.1) can be obtained as appropriate limit of theories of type I, II, II*.

We fix G and R as in (4.1). We denote respectively by V and E the set of vertices and edges of the graph represented by c_{vw} . We denote the number of vertices by $V = |V|$. The instanton partition function depends on the following quantities

- Ω -deformation parameters ϵ_1 and ϵ_2 , we set $\epsilon = \epsilon_1 + \epsilon_2$
- gauge couplings $q_v, v \in V$

- Coulomb parameters $a \in$ Cartan subalgebra of $\text{Lie}(G)$, $a_{\mathbf{v},\lambda}$ for $\mathbf{v} \in \mathcal{V}$, $\lambda = 1, \dots, N_{\mathbf{v}}$
- fundamental masses $m_{\mathbf{v},f}$, $\mathbf{v} \in \mathcal{V}$, $f = 1, \dots, n_{\mathbf{v}}$
- bifundamental masses⁷ $m_{\mathbf{e}}$, $\mathbf{e} \in \mathcal{E}$.

We package masses and Coulomb parameters into polynomials as follows

$$M_{\mathbf{v}}(x) = \prod_{f=1}^{n_{\mathbf{v}}} (x - m_{\mathbf{v},f}), \quad P_{\mathbf{v}}(x) = \prod_{\lambda=1}^{N_{\mathbf{v}}} (x - a_{\mathbf{v},\lambda}). \quad (4.2)$$

The contour integral representation of the Nekrasov partition function takes the form

$$\mathcal{Z} = \sum_{k_1, \dots, k_V}^{\infty} \left(\frac{\epsilon}{\epsilon_1 \epsilon_2} \right)^k \frac{q_1^{k_1}}{k_1!} \cdots \frac{q_V^{k_V}}{k_V!} \int \prod_{\mathbf{v} \in \mathcal{V}} \left([d\phi_{\mathbf{v}}] z_{\mathbf{v},k}^{\text{gauge}}(\phi_{\mathbf{v}}) z_{\mathbf{v},k}^{\text{fund}}(\phi_{\mathbf{v}}) \right) \prod_{\mathbf{e} \in \mathcal{E}} z_{\mathbf{e},k}^{\text{bifund}}(\phi_{s_{\mathbf{e}}}, \phi_{t_{\mathbf{e}}}) \quad (4.3)$$

where $k = \sum k_{\mathbf{v}}$ and $[d\phi_{\mathbf{v}}] = \prod_{I=1}^{k_{\mathbf{v}}} (2\pi i)^{-1} d\phi_{\mathbf{v},I}$. As above, \mathcal{V} and \mathcal{E} are respectively set of vertices and edges of the quiver. Moreover, given an edge \mathbf{e} , its orientation defines two vertices called $s_{\mathbf{e}}$ (source) and $t_{\mathbf{e}}$ (target). Using the expressions for $z_{\mathbf{v}}^{\text{gauge}}$, $z_{\mathbf{v}}^{\text{fund}}$ and $z_{\mathbf{e}}^{\text{bifund}}$ given in Appendix A.3 we rewrite the integrand of (4.3) as

$$\prod_{\mathbf{v} \in \mathcal{V}} \left[\left(\prod_{I=1}^{k_{\mathbf{v}}} Q_{\mathbf{v}}(\phi_{\mathbf{v},I}) \right) \left(\prod_{1 \leq I < J \leq k_{\mathbf{v}}} \mathcal{D}(\phi_{\mathbf{v},I} - \phi_{\mathbf{v},J}) \right) \right] \prod_{\mathbf{e} \in \mathcal{E}} \left[\prod_{I=1}^{k_{s_{\mathbf{e}}}} \prod_{J=1}^{k_{t_{\mathbf{e}}}} \mathcal{D}_{\mathbf{e}}(\phi_{s_{\mathbf{e}},I} - \phi_{t_{\mathbf{e}},J}) \right] \quad (4.4)$$

where

$$Q_{\mathbf{v}}(x) = \frac{M_{\mathbf{v}}(x)}{P_{\mathbf{v}}(x) P_{\mathbf{v}}(x + \epsilon)} \prod_{\mathbf{v} \xrightarrow{\mathbf{e}} \mathbf{v}'} P_{\mathbf{v}'}(x - m_{\mathbf{e}}) \prod_{\mathbf{v} \xleftarrow{\mathbf{e}} \mathbf{v}'} P_{\mathbf{v}'}(x + m_{\mathbf{e}} + \epsilon), \quad (4.5)$$

$$\mathcal{D}(x) = \Delta(x) \Delta(-x), \quad \mathcal{D}_{\mathbf{e}}(x) = [\Delta(m_{\mathbf{e}} + x) \Delta(m_{\mathbf{e}} - x)]^{-1}, \quad \Delta(x) = \frac{x(x + \epsilon_1 + \epsilon_2)}{(x + \epsilon_1)(x + \epsilon_2)}, \quad (4.6)$$

Note that \mathcal{D} is the same as for pure SYM, see (1.5). Moreover, we have the manifest symmetry $\mathcal{D}(x) = \mathcal{D}(-x)$ and $\mathcal{D}_{\mathbf{e}}(x) = \mathcal{D}_{\mathbf{e}}(-x)$.

Explanations regarding the origin of this expression are collected in Appendix A. The expression for the contribution to the partition function from bifundamental matter can be found in [57] and [58]. For more general gauge groups and representation see e.g. [59, 54] and references therein.

⁷To avoid confusion we recall that in our conventions a bifundamental hypermultiplet associated to the edge \mathbf{e} corresponds to a factor $(\square_{s_{\mathbf{e}}}, \overline{\square}_{t_{\mathbf{e}}}) + (\square_{t_{\mathbf{e}}}, \overline{\square}_{s_{\mathbf{e}}})$ in (4.1).

4.1 TBA for quiver gauge theories

The partition function (4.3) in the limit $\epsilon_2 \rightarrow 0$ can be studied in a similar way as for the simpler case of pure $SU(N)$ SYM, see (1.4). In particular it can be characterized as solution of certain TBA equations given below. The important observation to be made is that as long as the bifundamental masses m_e are large compared to ϵ_2 , the pair-wise interaction term in the non-ideal gas interpretation of (4.4) splits into two terms with different scales, compare to (2.2). The first term $\frac{x^2}{x^2 - \epsilon_2^2}$, present for each gauge group, gives rise to the dilogarithm function in the TBA. The remaining factor in the pair-wise interaction is responsible for the kernel term in the TBA non-linear integral equation.

The resulting TBA is summarized as follows. As in [18], the partition function can be written as the critical value of the Yang-Yang functional

$$\mathcal{Y}(\rho, \varphi) = \frac{1}{2} \sum_{v,w} \int \frac{dx}{2\pi i} \frac{dy}{2\pi i} \rho_v(x) G_{vw}(x-y) \rho_w(y) + \sum_v \int \frac{dx}{2\pi i} \left[\rho_v(x) \varphi_v(x) + \text{Li}_2(q_v Q_r(x) e^{-\varphi_v(x)}) \right], \quad (4.7)$$

where the variations with respect to $\varphi_v(x)$ and $\rho_v(x)$ give

$$\rho_v(x) = -\log [1 - q_v Q_v(x) e^{-\varphi_v(x)}], \quad \varphi_v(x) = -\sum_w \int \frac{dy}{2\pi i} G_{vw}(x-y) \rho_w(y). \quad (4.8)$$

The twisted superpotential for the quiver gauge theory is obtained as

$$\mathcal{W} = \sum_v \int \frac{dx}{2\pi i} \left[-\frac{1}{2} \varphi_v(x) \log (1 - q_v Q_v(x) e^{-\varphi_v(x)}) + \text{Li}_2(q_v Q_v(x) e^{-\varphi_v(x)}) \right], \quad (4.9)$$

where $\varphi_v(x)$ satisfy the TBA equation

$$\varphi_v(x) = \sum_w \int \frac{dy}{2\pi i} G_{vw}(x-y) \log [1 - q_w Q_w(y) e^{-\varphi_w(y)}]. \quad (4.10)$$

The various functions which contain the data of quiver theory are

$$G_{vw}(x) = \frac{d}{dx} \left[\delta_{vw} \log \left(\frac{x + \epsilon_1}{x - \epsilon_1} \right) + c_{vw} \log \left(\frac{x + m_{vw}}{x + m_{vw} + \epsilon_1} \right) + c_{wv} \log \left(\frac{x - m_{vw} - \epsilon_1}{x - m_{vw}} \right) \right], \quad (4.11)$$

and $Q_v(x)$ defined in (4.5). In (4.11) we defined $m_{vw} := m_e$ if the vertices v and w are connected by the edge e and $m_{vw} := 0$ otherwise. The expression for the propagator $G_{vw}(x)$ given above is obtained as as

$$G_{vw}(x) = \begin{cases} \text{Limit}_{\epsilon_2 \rightarrow 0} \frac{\mathcal{D}(x)-1}{\epsilon_2} & \text{for } v = w \\ \text{Limit}_{\epsilon_2 \rightarrow 0} \frac{\mathcal{D}_e(x)-1}{\epsilon_2} & \text{for } v \xrightarrow{e} w \\ 0 & \text{otherwise} \end{cases} \quad (4.12)$$

Note that in this definition we have $G_{vw}(x) = G_{wv}(-x)$.

5 Conclusion and Discussion

In this paper we studied the instanton partition functions of four dimensional $\mathcal{N} = 2$ gauge theories in a special limit of Ω deformation parameters, namely, taking $\epsilon_2 \rightarrow 0$ but keeping ϵ_1 finite. We show explicitly that the instanton part of the twisted superpotential \mathcal{W} , which in statistical mechanics language is equal to the free energy, satisfies TBA equations. It is thus naturally identified with the Yang-Yang function of some quantum integrable model [18]. Based on the proof, we are also able to generalize the correspondence to a general class of $\mathcal{N} = 2$ quiver theories. In this case one obtains a set of coupled TBA equations with one equation for each node of the quiver.

The starting point of the derivation is the contour integral representation of the Nekrasov partition function. It can be interpreted as the grand canonical partition function of a non-ideal gas of particles. Due to the singular structure of the two-particles interaction, an interesting effective description of this gas emerges in the $\epsilon_2 \rightarrow 0$ limit. In this effective description the “new particles” are clusters of the original particles and are subject to certain tree-level interaction only. This picture is best obtained by employing the so-called iterated Mayer expansion [50]. This structure is nicely captured by the TBA, in particular the clustering of particles and their effective interaction is nicely reproduced by the presence of the Li_2 function⁸. An all order proof in the instanton counting parameter q of the TBA equations is given applying a $(0+1)$ -dimensional path integral trick, dating back to Polyakov [51].

In studying the NS limit of the instanton partition function, one has to face the problem of evaluating integrals when some parameter, in our case ϵ_2 , is small. There is a systematic way to do so, particularly successful for calculating Feynman integrals, called the method of expansion by regions. While our results are obtained also independently of this method, it provides a simple way to single out the leading $\frac{1}{\epsilon_2}$ behavior of the logarithm of the partition function. This structure is somewhat reminiscent of the exponentiation of infrared divergences in gauge theories. It would be interesting to further study the applicability of this method to the type of integral considered here.

In this paper we considered quiver gauge theories involving only $U(N)$ gauge group factors and matter in the fundamental/antifundamental or bifundamental representations. It is natural to extend this work by deriving TBA equation in the case of the other classical gauge groups $SO(N)$, $SP(2N)$ and to other matter content, e.g. symmetric and antisymmetric representations of $U(N)$ or trifundamental of $U(2)$ ³. Other interesting generalizations are the deformation of the partition function by chiral ring operators and the five dimensional lift of the partition function. The details of these studies will be reported elsewhere.

One of the most challenging question for the future is to understand how the proposed TBA equations emerge from the quantum integrable model point of view. In the case of

⁸This makes it drastically different from some other TBA obtained by simply applying saddle point method [60, 61, 62]. See also [63].

pure $SU(N)$ $\mathcal{N} = 2$ super Yang-Mills, which corresponds to the periodic Toda chain with N sites, this has been achieved in [24]. In the derivation a crucial role is played by the so-called Baxter Q-operator, which is currently the most powerful and universal tool for determining the spectrum of quantum integrable systems. Despite successful applications in many cases, see [64] (and also [65]) for recent progress and relevant references, a complete theory of Q-operators is still to be established. In particular a systematic construction is currently not available in the case for which the representation of the relevant quantum group is not of highest-weight type, see [66] and references therein for some progress in this direction. Via the NS proposal, quantum integrable systems of this type can be systematically solved by gauge theory methods, offering an entirely new perspective on the structure underlying integrable models.

TBA equations of a similar type also appeared in the context of wall crossing effects for the BPS spectrum in $\mathcal{N} = 2$ gauge theories [67, 68]. It would be interesting to study the connection with our work and find the quantum integrable models corresponding to the TBA appearing there. Another connection is to the study of minimal surfaces in AdS [69, 70]. In this case, the area of such surfaces is equal, via the AdS/CFT correspondence, to the strong coupling limit of the logarithm of null Wilson-loop expectation values in $\mathcal{N} = 4$ SYM. In [71, 72], it has been proposed that the exact expectation value can be written as a sum of multiple integrals involving some basic building blocks called pentagon transition functions. This form is structurally identical to the contour integral representation of the Nekrasov partition function. Moreover, the strong coupling limit correspond to the NS limit with $\frac{1}{\sqrt{\lambda}} \sim \epsilon_2$. Techniques of the type presented in this paper can be also used to show, directly from this representation of the Wilson loop, that in the strong coupling limit it satisfies the TBA equations derived from the analysis of classical strings in AdS [73, 74]. It would be very interesting to study the possible connection with our work in this respect.

Acknowledgements

We are greatly indebted to Jörg Teschner for suggesting this problem and valuable discussions at various stages of the project. We would like to thank Yasuyuki Hatsuda and Elli Pomoni for very useful discussions and collaboration on related topics. We would also like to thank Marcos Mariño for his interest in this work and encouragement. An important part of the work was done when G.Yang was a postdoc in the University of Hamburg where he was supported by the German Science Foundation (DFG) within the Collaborative Research Center 676 “Particles, Strings and the Early Universe”. C. Meneghelli is partially supported by a DFG grant in the framework of the SFB 676 “Particles, Strings, and the Early Universe”. G.Yang is supported by a DFG grant in the framework of the SFB 647 “Space-Time-Matter”.

A On the integral representation form of the Nekrasov partition function

In this Appendix we present a brief, non-self-consistent review, together with a collection of references, on the origin of the contour integral form of the instanton partition function (1.4), [8, 10].

The first step is to localize the path integral for the given $\mathcal{N} = 2$ supersymmetric gauge theory in the Ω -background to configuration of fields satisfying the self-duality equation $F_+(A) = 0$ and the Dirac equation $D_A q = 0$, see e.g. [75]. Solutions to the self-duality equation $F_+(A) = 0$ modulo gauge transformations that are trivial at infinity defines the instanton moduli space $\mathfrak{M}^{\text{inst}}$. The space of solutions to the Dirac equation forms a fiber over the instanton moduli space.

If the gauge group is one of the classical groups $U(N)$, $SO(N)$, $SP(2N)$ (or a product thereof), $\mathfrak{M}^{\text{inst}}$ admits a particular nice description, the ADHM construction [76]. Let us review this construction for gauge group $U(N)$ and instanton number k . First introduce the linear data

$$x = (\mathbf{B}_1, \mathbf{B}_2, \mathbf{I}, \mathbf{J}) \in \mathcal{X}_{N,k} := \text{Hom}(V, V) \oplus \text{Hom}(V, V) \oplus \text{Hom}(W, V) \oplus \text{Hom}(V, W), \quad (\text{A.1})$$

where $V \simeq \mathbb{C}^k$ and $W \simeq \mathbb{C}^N$ carry the defining $GL(k)$ and $GL(N)$ action. We refer to (A.1) as ADHM matrices. On such matrices the group $GL(2) \times GL(k) \times GL(N)$ acts as follows

$$(\mathbf{B}_\alpha, \mathbf{I}, \mathbf{J}) \mapsto (g_\phi (M_\alpha^\beta \mathbf{B}_\beta) g_\phi^{-1}, g_a \mathbf{I} g_\phi^{-1}, \det(M) g_\phi \mathbf{J} g_a^{-1}), \quad (\text{A.2})$$

where $\alpha, \beta = 1, 2$, the summation over β is understood and

$$M \in GL(2), \quad g_a \in GL(N), \quad g_\phi \in GL(k). \quad (\text{A.3})$$

The next step is to introduce the ADHM equations

$$\mu_{\mathbb{R}} := \mathbf{I} \mathbf{I}^\dagger - \mathbf{J}^\dagger \mathbf{J} + [\mathbf{B}_1, \mathbf{B}_1^\dagger] + [\mathbf{B}_2, \mathbf{B}_2^\dagger] = 0, \quad (\text{A.4})$$

$$\mu_{\mathbb{C}} := \mathbf{I} \mathbf{J} + [\mathbf{B}_1, \mathbf{B}_2] = 0. \quad (\text{A.5})$$

These equations are covariant under the group action given by (A.2). The k -instanton moduli space is described as

$$\mathfrak{M}_{N,k} = \{x \in \mathcal{X}_{N,k} \text{ such that } \mu_{\mathbb{C}}(x) = 0\} / GL(k). \quad (\text{A.6})$$

There are two non-trivial steps involved in this definition: (1) impose the ADHM equations, (2) divide the resulting space by the action of $GL(k)$. In the following we will see how these two steps give rise to the contour integral form of the instanton partition function following two different procedures. We anticipate that in both cases the integration over (ϕ_1, \dots, ϕ_k) in (1.4) originates from quotienting over $GL(k)$.

A.1 D-instanton action and its localization

There is a natural way to understand the ADHM construction within the language of branes [77, 79, 78]. In the simplest example, the starting point is a system of k D(-1) branes and N D3-branes. The idea is to look at this system either as the theory leaving on the D3-branes, which corresponds to a four dimensional $\mathcal{N} = 2$ gauge theory, or as the $d = 0+0$ dimensional theory on the D(-1) brane. In the latter case one identifies the ADHM equations with D and F-flatness conditions of the auxiliary 0-dimensional theory. In this way the instanton moduli space $\mathfrak{M}_{N,k}$ is identified with the Higgs branch of the theory on the D-instanton.

From the point of view of the four dimensional gauge theory the D-instanton partition function originates as an integral over the k-instanton moduli space. Following [8], see also [75], one can present the integration over the instanton moduli space $\mathfrak{M}_{N,k}$, see (A.6), in terms of the one over $\mathcal{X}_{N,k}$, see (A.1), and certain ‘‘auxiliary fields’’. The latter, as emphasized after equation (A.6), are introduced in order to implement the ADHM constraints and the operation of taking quotient by $GL(k)$. See also [83, 84].

The corresponding partition function reads

$$\mathcal{Z}_k = \int \frac{\mathcal{D}\phi}{\text{Vol}(U(k))} [\mathcal{D}x \mathcal{D}\psi] [\mathcal{D}\chi \mathcal{D}H] [\mathcal{D}\bar{\phi} \mathcal{D}\eta] e^{-S_\Omega}. \quad (\text{A.7})$$

This equation needs some explanation. Measure factors in [...] correspond to BRST doublets

$$\mathcal{Q} \begin{pmatrix} x \\ \chi \\ \bar{\phi} \end{pmatrix} = \begin{pmatrix} \psi \\ H \\ \eta \end{pmatrix}, \quad \mathcal{Q} \begin{pmatrix} \psi \\ H \\ \eta \end{pmatrix} = \delta_{\phi,a,\epsilon_{1,2}}^{\text{torus}} \begin{pmatrix} x \\ \chi \\ \bar{\phi} \end{pmatrix}. \quad (\text{A.8})$$

The variable x corresponds to ADHM matrixes (A.1), $\chi = (\chi_{\mathbb{R}}, \chi_{\mathbb{C}})$ belong to the same space as the ADHM equations (A.4), (A.5) and $\bar{\phi} \in \text{Hom}(V, V)$. Their BRST partners (ψ, H, η) have the same transformation properties under $GL(N) \times GL(k) \times GL(2)$. The variable ϕ has no BRST partner, it satisfies $\mathcal{Q}\phi = 0$. $(x, H, \bar{\phi}, \phi)$ are bosonic, (ψ, χ, η) fermionic. The action of $\delta^{\text{torus}} \in \text{Lie}(\text{torus})$ is defined as taking the infinitesimal version of the transformation (A.2) with

$$g_a - 1 \sim \text{diag}(a_1, \dots, a_N), \quad g_\phi - 1 \sim \text{diag}(\phi_1, \dots, \phi_k), \quad M - 1 \sim \text{diag}(\epsilon_1, \epsilon_2). \quad (\text{A.9})$$

It is clear from the action of \mathcal{Q} that $\mathcal{Q}^2 = \delta^{\text{torus}}$. The action has the schematic form

$$S_\Omega = \mathcal{Q} (\chi \cdot \mu(x) + t \chi \cdot H + t' \eta [\phi, \bar{\phi}] + t'' \psi \bar{\phi} x), \quad (\text{A.10})$$

where $\mu(x)$ are the ADHM equations. The parameter t, t', t'' have a fixed value in the original problem, but, as the partition function is independent of their values, we can calculate it for the most convenient choice. The suffix Ω refers to the so-called Ω -deformation introduced to regularize the infinite volume of moduli space [8, 9].

The variables $(\eta, \chi, H, \bar{\phi})$ come into a quartet that can be integrated out in the following way [8, 10, 75]. Add to the action a Q-exact term

$$\delta S = s \mathcal{Q}(\chi_{\mathbb{R}} \cdot \bar{\phi}) + s' \mathcal{Q}(x \psi). \quad (\text{A.11})$$

For $s \rightarrow \infty$, the $H_{\mathbb{R}}, \chi_{\mathbb{R}}$ integration produces the constraints $\bar{\phi} = 0 = \eta$. Next taking $t \rightarrow \infty$ the $H_{\mathbb{C}}$ integration is Gaussian and produces a trivial factor, the $\chi_{\mathbb{C}}$ integration is Gaussian (fermionic) and produces the factor

$$\det_{\mu_{\mathbb{C}}} \mathcal{Q}^2 = \epsilon^k \prod_{1 \leq I < J \leq k} (\phi_{IJ}^2 - \epsilon^2). \quad (\text{A.12})$$

The notation $\det_{\mu_{\mathbb{C}}}$ refers to taking the determinant in the vector space to which $\chi_{\mathbb{C}}$ belongs which is the same as the space of ADHM equations $\mu_{\mathbb{C}}$, see (A.5). To calculate the determinant we also reduce the ϕ integration from $GL(k)$ to its maximal torus

$$\mathcal{D}\phi \rightarrow \frac{1}{k!} \prod_{I=1}^k \frac{d\phi_I}{2\pi i} \prod_{1 \leq I < J \leq k} \phi_{IJ}^2. \quad (\text{A.13})$$

Finally the integration over (x, ψ) , is done sending $s' \rightarrow \infty$. The $(\mathbf{B}_{1,2}, \psi_{B_{1,2}})$ integration produces the factor

$$\frac{1}{\det_{\mathbf{B}_{1,2}} \mathcal{Q}^2} = \frac{1}{\epsilon_{1,2}^k} \prod_{1 \leq I < J \leq k} \frac{1}{\phi_{IJ}^2 - \epsilon_{1,2}^2}, \quad (\text{A.14})$$

and the $(\mathbf{I}, \psi_I), (\mathbf{J}, \psi_J)$ integration produces the factor

$$\frac{1}{\det_{\mathbf{I}, \mathbf{J}} \mathcal{Q}^2} = \prod_{I=1}^k \prod_{\lambda=1}^N \frac{1}{(\phi_I - a_{\lambda})(\phi_I - a_{\lambda} + \epsilon)}. \quad (\text{A.15})$$

We have thus performed all the integration in (A.7) except for (ϕ_1, \dots, ϕ_k) . Putting together (A.12), (A.13), (A.14), (A.15) one recognizes that \mathcal{Z}_k given in (A.7) coincides with (1.4). This ends the discussion of pure SYM with gauge group $U(N)$. The discussion can be generalized to other classical gauge groups and quiver gauge theories, see e.g. [58] and references therein.

A.2 The instanton partition function from 5d perspective

As a different picture, we consider the lift of the given $\mathcal{N} = 2$ four dimensional gauge theory to five dimensions. The Ω deformation can be understood as imposing the identification $(x, y) \sim (e^{\beta\Omega}x, y + \beta)$, where y is the coordinate of the circle of compactification and Ω the generator of an infinitesimal $SO(4)$ rotation. The important fact is that the 5d partition function $\mathcal{Z}_k^{d=5}$ reduces to the twisted Witten index for the supersymmetric quantum mechanics on the instanton moduli space $\mathfrak{M}_{N,k}$ [12], as fermions have periodic boundary condition (up to the twist) along y direction.

As described in details in [80], by the ADHM description of the instanton moduli space (A.6), the 5d index can be given as

$$\mathcal{Z}_k^{d=5} = \frac{1}{|W_k| \text{Vol}(\mathbb{T}_k)} \int_{\mathbb{T}_k} \prod_{I=1}^k d\phi_I \prod_{\alpha \in \Delta_+} (e^{i\langle \alpha, \phi \rangle} - e^{-i\langle \alpha, \phi \rangle}) \frac{\prod(1 - \varphi_{\text{ADHM equations}})}{\prod(1 - \omega_{\text{ADHM matrices}})}. \quad (\text{A.16})$$

The factors $\varphi_{\text{ADHM equations}}$ and $\omega_{\text{ADHM matrices}}$ are the weights under a certain torus action, of which (A.9) is the infinitesimal version, of the ADHM equations (A.5) and ADHM matrices (A.1) respectively. The structure of (A.16) originates from the two step procedure emphasized after equation (A.6): the term $\prod(1 - \varphi_{\text{ADHM equations}})$ encodes the constraints from the ADHM equation, the integration, and the Vandermonde factor, implements the $GL(k)$ quotient. In the following we spell out this formula for pure $\mathcal{N} = 2$ SYM with gauge group $U(N)$.

Consider an element in the maximal torus $\mathbb{T} \subset GL(k) \times GL(N) \times GL(2)$ given by $g_\phi = \text{diag}(e^{i\beta\phi_1}, \dots, e^{i\beta\phi_k})$, $g_a = \text{diag}(e^{i\beta a_1}, \dots, e^{i\beta a_N})$, $M = \text{diag}(q_1, q_2)$ with $|q_1|, |q_2| < 1$, compare to (A.2). The eigenvalues of the torus action on the ADHM matrices $\mathbf{B}_\alpha, \mathbf{I}, \mathbf{J}$ are

$$q_\alpha e^{i\beta\phi_{IJ}}, \quad e^{i\beta(a_\lambda - \phi_I)}, \quad q_1 q_2 e^{i\beta(\phi_I - a_\lambda)}, \quad (\text{A.17})$$

where $\alpha = 1, 2$, $I, J = 1, \dots, k$, and $\lambda = 1, \dots, N$. One thus obtains

$$\frac{1}{\prod(1 - \omega_{\text{ADHM matrices}})} \Big|_{\mathbf{I}, \mathbf{J}} = \prod_{I=1}^k \frac{1}{P_{+\beta}(e^{-i\beta\phi_I}) P_{-\beta}(q_1 q_2 e^{i\beta\phi_I})}, \quad (\text{A.18})$$

$$\frac{1}{\prod(1 - \omega_{\text{ADHM matrices}})} \Big|_{\mathbf{B}_{1,2}} = \frac{1}{(1 - q_{1,2})^k} \frac{1}{\Delta_\beta(q_{1,2})}, \quad (\text{A.19})$$

where

$$P_\beta(t) := \prod_{\lambda=1}^N (1 - t e^{i\beta a_\lambda}), \quad \Delta_\beta(t) := \prod_{I \neq J} (1 - t e^{i\beta\phi_{IJ}}). \quad (\text{A.20})$$

It is not a coincidence that these equations are in correspondence with (A.15), (A.14).

Next consider the contribution from the *complex* ADHM equations (A.5). Its weight under the thorus action gives

$$\prod(1 - \varphi_{\text{ADHM equations}}) = (1 - q_1 q_2)^k \Delta_\beta(q_1 q_2). \quad (\text{A.21})$$

Compare to (A.12). Finally the Vandermonde is equal to $\Delta_\beta(1)$. Putting things together according to equation (A.16) we obtain

$$\frac{(1 - q_1 q_2)^k}{(1 - q_1)^k (1 - q_2)^k} \frac{\Delta_\beta(1) \Delta_\beta(q_1 q_2)}{\Delta_\beta(q_1) \Delta_\beta(q_2)} \prod_{I=1}^k \frac{1}{P_\beta(e^{-i\beta\phi_I}) P_\beta(q_1 q_2 e^{-i\beta\phi_I})}. \quad (\text{A.22})$$

The cases of $SO(N)$ and $SP(2N)$ are similar upon modifying ADHM data and related thorus action, [80], see also [81]. Matter can also be included in this picture. The limit $\beta \rightarrow 0$ can be easily taken upon setting $q_{1,2} = e^{i\beta\epsilon_{1,2}}$. Notice that the condition $|q_1|, |q_2| < 1$ translates into a small imaginary part for $\epsilon_{1,2}$.

A.3 Building blocks for quiver gauge theories

We now present the recipe to obtain the integrand for the instanton partition function corresponding to the class of quiver gauge theories introduced in Section 4. To which we refer for the notation used here. From the discussion above it is expected that the integrand should be written in terms of the following data:

- (1) an element $q = q(\epsilon_{1,2}, a, \phi, m) \in \mathbb{T} \subset SO(4)_{\text{Lorentz}} \times G_{\text{gauge}} \times G_{\text{inst}} \times G_{\text{flavour}}$,
- (2) a ‘‘collection’’ of weights ω_ℓ , more precisely eigenvalues $e^{i\omega_\ell}$, under this torus action.

From this data the integrand can be constructed using the following rule

$$\sum_\ell n_\ell e^{i\omega_\ell} \mapsto \prod_\ell (\omega_\ell)^{-n_\ell}, \quad (\text{A.23})$$

where $n_\ell = \pm 1$.

Let us introduce the following modules for the torus action:

$$W_v \simeq \mathbb{C}^{N_v}, \quad V_v \simeq \mathbb{C}^{k_v}, \quad M_v \simeq \mathbb{C}^{n_v}, \quad L \simeq \mathbb{C}^2 \simeq \mathbb{R}^4. \quad (\text{A.24})$$

They carry, respectively, the defining representation of $GL(N_v)$, $GL(k_v)$, $GL(n_v)$, $SO(4)_{\text{Lorentz}}$ and thus of the maximal torus. We denote the corresponding characters as $\chi_{W,V,M} = \text{tr}_{W,V,M}(q)$, for example $\chi_V = \sum_{I=1}^k e^{i\phi_I}$. We denote their conjugate modules as \overline{W}_v , \overline{V}_v , \overline{M}_v .

From the description of the tangent space to the instanton moduli space in terms of linearized ADHM equations and infinitesimal gauge transformations, see [11], one finds, for each gauge group factor

$$\text{Ch}(T\mathfrak{M}_{N,k}) = \chi_{\overline{W}} \chi_V - (e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1) \chi_V \chi_{\overline{V}} + e^{i\epsilon} \chi_{\overline{V}} \chi_W. \quad (\text{A.25})$$

Appling the rule (A.23) to this character one obtains the integrand (1.4). The contribution of bifundamental matter is given by

$$\text{Ch}(\square_v, \overline{\square}_w) = -e^{i m_{vw}} [\chi_{W_v} \chi_{\overline{V}_w} - (e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1) \chi_{V_v} \chi_{\overline{V}_w} + e^{i\epsilon} \chi_{V_v} \chi_{\overline{W}_w}]. \quad (\text{A.26})$$

Notice the overall minus sign compared to (A.25). Applying the rule (A.23) one obtains

$$\text{Ch}(\square_v, \overline{\square}_w) \mapsto (-1)^{k_w N_v} \prod_{I=1}^{k_v} \prod_{J=1}^{k_w} \frac{P_v(\phi_{w,J} - m_{vw}) P_w(\phi_{v,I} + m_{vw} + \epsilon)}{\Delta(m_{vw} + \phi_{v,I} - \phi_{w,J})} =: \tilde{z}_{k_v k_w}^{vw}. \quad (\text{A.27})$$

From this expression it follows that the contribution from $\text{Ch}(\square_v, \overline{\square}_w) + \text{Ch}(\square_w, \overline{\square}_v)$ is given by

$$z_k^{\text{bifund}} = \tilde{z}_{k_v k_w}^{vw} \tilde{z}_{k_w k_v}^{wv}. \quad (\text{A.28})$$

Finally fundamental matter gives the character

$$\text{Ch}(\square_v) = -\chi_V \chi_{\overline{M}} \xrightarrow{\text{(A.23)}} z_k^{\text{fund}} = \prod_{I=1}^k \prod_{f=1}^n (\phi_I - m_f) . \quad (\text{A.29})$$

The expressions above can also be found in [57, 58].

A.4 Contour and classification of poles

The poles contributing to (1.4) correspond to solutions of simple algebraic equations. As we will now review following [85], these solutions are classified by N-tuple of Young diagrams [11]. Let

$$A(x) := \prod_{\lambda=1}^N (x - a_\lambda), \quad E(x) := \frac{(x - \epsilon_1)(x - \epsilon_2)}{x(x - \epsilon_1 - \epsilon_2)} . \quad (\text{A.30})$$

We want to classify solutions of

$$A(x_i) \prod_{j \neq i} E(x_i - x_j), \quad \text{for } i = 1, \dots, k, \quad (\text{A.31})$$

modulo permutations of $\{x_i\}$. Notice that, compare to (1.4), only ‘‘half of the poles’’ of the integrand contributes. This is so as a consequence of the choice of contour. It is clear that at least one x has to be a zero of A . Without loss of generality we set $x_1 = a_{\lambda_1}$ and get $k-1$ equations

$$A'(x_i) \prod_{\substack{j > 1 \\ j \neq i}} E(x_i - x_j), \quad \text{for } i = 2, \dots, k, \quad (\text{A.32})$$

where $A'(x) := A(x)E(x - a_{\lambda_1})$. For the same argument as above applied now to (A.32), we set $A'(x_2) = 0$. We have to recall that we must keep $x_i - x_j \neq 0$ and $x_i - x_j \neq \epsilon_1 + \epsilon_2$ otherwise the denominator in $E(x)$ vanishes. Iterating this procedure one finds that solutions of (A.31) modulo permutations are classified by the N-tuple of Young diagrams. For example for $k = 3$ one has

$$\{a_{\lambda_1}, a_{\lambda_2}, a_{\lambda_3}\}, \quad \{a_{\lambda_1}, a_{\lambda_2}, a_{\lambda_2} + \epsilon_\alpha\}, \quad \{a_{\lambda_1}, a_{\lambda_1} + \epsilon_\alpha, a_{\lambda_1} + 2\epsilon_\alpha\}, \quad \{a_{\lambda_1}, a_{\lambda_1} + \epsilon_1, a_{\lambda_1} + \epsilon_2\}, \quad (\text{A.33})$$

where $\alpha = 1, 2$ and the a_λ are different from each other.

B Useful formulas

B.1 Computing (2.16)

In the following we present the explicit calculation of the residue corresponding to the pole (2.15). First we introduce the notation

$$\mathbf{Y}_1 = \{1, \dots, s_1\}, \quad \mathbf{Y}_2 = \{s_1, \dots, s_1 + s_2\}, \dots, \mathbf{Y}_N = \{s_1 + \dots + s_{N-1}, \dots, k\}. \quad (\text{B.1})$$

The next step is to split the interaction term $\prod \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2}$ in the integrand as

$$\prod_{1 \leq I < J \leq k} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} = \left[\prod_{\lambda=1}^N \prod_{\substack{I < J \\ I, J \in \mathbf{Y}_\lambda}} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} \right] \prod_{\lambda < \lambda'} \prod_{I_\lambda, I_{\lambda'}} \frac{\phi_{I_\lambda I_{\lambda'}}^2}{\phi_{I_\lambda I_{\lambda'}}^2 - \epsilon_2^2}, \quad (\text{B.2})$$

Only the factor in square bracket on the right hand side is singular in the limit in which ϕ_I take the value give by (2.15). For these terms we use the simple result

$$\lim_{\phi_{\kappa_\lambda+I_\lambda} \rightarrow a_\lambda + (I_\lambda - 1)\epsilon_2} \prod_{I=2}^{s_\lambda} (\phi_{\kappa_\lambda+I_\lambda} - a_\lambda - (I_\lambda - 1)\epsilon_2) \prod_{\substack{I < J \\ I, J \in \mathbf{Y}_\lambda}} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} = \frac{\epsilon_2^{s_\lambda - 1}}{s_\lambda}, \quad (\text{B.3})$$

We also use

$$\lim_{\phi_{\kappa_\lambda+1} \rightarrow a_\lambda} (\phi_{\kappa_\lambda+1} - a_\lambda) \prod_{I \in \mathbf{Y}_\lambda} Q(\phi_I) = \frac{1}{\epsilon_2^{s_\lambda - 1}} \frac{1}{(s_\lambda - 1)!} \prod_{I=1}^{s_\lambda} \frac{Q^{(\lambda)}(a_\lambda + (I - 1)\epsilon_2)}{\epsilon_1 + I\epsilon_2}. \quad (\text{B.4})$$

where

$$Q^{(\lambda_1)}(x) = \frac{1}{P^{(\lambda_1)}(x)P^{(\lambda_1)}(x + \epsilon)}, \quad P^{(\lambda_1)}(x) = \prod_{\lambda \neq \lambda_1} (x - \lambda). \quad (\text{B.5})$$

The result (2.16) immediately follows.

B.2 Exponentiation of residues

One has

$$\sum_{k=0}^{\infty} \frac{q^k}{k! \epsilon^k} \prod_{I=1}^k f(x + (I - 1)\epsilon) = \exp \left(\frac{1}{\epsilon} \sum_{\ell=1}^{\infty} \frac{q^\ell}{\ell^2} \left(\frac{1}{(\ell - 1)!} \frac{d^{\ell-1}}{dx^{\ell-1}} f(x)^\ell \right) + \mathcal{O}(\epsilon^0) \right) \quad (\text{B.6})$$

as formal series in ϵ .

We present the following generalization of (B.6)

$$\begin{aligned} & \sum_{s_1, \dots, s_N \geq 1} \prod_{\lambda=1}^N \left(\frac{q^{s_\lambda} \prod_{I=1}^{s_\lambda} A(x_\lambda + (I-1)\epsilon))}{s_\lambda! \epsilon^{s_\lambda}} \right) \prod_{1 \leq \lambda < \lambda' \leq N} \prod_{I_\lambda=1}^{s_\lambda} \prod_{I_{\lambda'}=1}^{s_{\lambda'}} \frac{(x_{\lambda\lambda'} + (I_\lambda - I_{\lambda'})\epsilon)^2}{(x_{\lambda\lambda'} + (I_\lambda - I_{\lambda'})\epsilon)^2 - \epsilon^2} \\ & = \exp \left(\frac{1}{\epsilon} \sum_{\lambda=1}^N \sum_{\ell=1}^{\infty} \frac{q^\ell}{\ell^2} \left(\frac{1}{(\ell-1)!} \frac{d^{\ell-1}}{dx_\lambda^{\ell-1}} A(x_\lambda)^\ell \right) + \mathcal{O}(\epsilon^0) \right) \end{aligned} \quad (\text{B.7})$$

as formal series in ϵ . We do not present a proof of this result, but it can be easily checked order by order in q .

B.3 Illustrative Examples for Iterated Mayer expansion

In the following we spell out the definition (2.36) for $k = 2, 3$. For $k = 2$ (2.36) reads

$$\mathfrak{S}_{\{1,2\}} = S_{\{1,2\}} \mathfrak{g}_{\{1,2\}} = (\mathcal{D}_{12} - 1) \tilde{\mathcal{D}}_{12}, \quad \mathfrak{S}_{\{1\},\{2\}} = \mathfrak{f}_{\{1\},\{2\}} = \tilde{\mathcal{D}}_{12} - 1, \quad (\text{B.8})$$

and the sum of this two terms is by the definitions above \mathcal{F}_{12} . For $k = 3$, (2.36) reads

$$\mathfrak{S}_{\{1,2,3\}} = S_{\{1,2,3\}} \mathfrak{g}_{\{1,2,3\}}, \quad (\text{B.9})$$

$$\mathfrak{S}_{\{I,J\},\{K\}} = S_{\{I,J\}} \mathfrak{g}_{\{I,J\}} \mathfrak{f}_{\{I,J\},\{K\}}, \quad (\text{B.10})$$

$$\begin{aligned} \mathfrak{S}_{\{1\},\{2\},\{3\}} = & \mathfrak{f}_{\{1\},\{2\}} \mathfrak{f}_{\{2\},\{3\}} + \mathfrak{f}_{\{1\},\{3\}} \mathfrak{f}_{\{2\},\{3\}} + \mathfrak{f}_{\{1\},\{2\}} \mathfrak{f}_{\{1\},\{3\}} \\ & + \mathfrak{f}_{\{1\},\{2\}} \mathfrak{f}_{\{2\},\{3\}} \mathfrak{f}_{\{1\},\{3\}}. \end{aligned} \quad (\text{B.11})$$

C Expansion by regions for instanton partition functions

One recurrent question in this paper is how to evaluate multiple integrals as a Laurent expansion in some small parameter⁹, namely ϵ_2 . There is a powerful method to address this question, it is the so called method of *expansion by regions*, usually applied to the evaluation of Feynman integrals [49, 86]. The method can briefly summarized as follows: (1) divide the integration domain into regions and expand the integrand in a Taylor series in small parameters, (2) extend the integration to the full domain of integration, (3) set the scaleless integrals to zero. It has been applied successfully in many situations in the context of Feynman integrals, see [87] for a recent study including proof under some assumptions. We apply the same idea to the integrals of instanton partition functions.

It might look surprising that the prescription above works at all. In particular one may worry about multiple counting of contributions. Let us describe how this problem is

⁹We are actually interested only in the leading $\frac{1}{\epsilon_2}$ behavior of the logarithm of the partition function.

circumvented in a simple example, the $k = 2$ case in (2.10)

$$F = \int \frac{d\phi_1}{2\pi i} \frac{d\phi_2}{2\pi i} Q(\phi_1)Q(\phi_2) \frac{\phi_{12}^2}{\phi_{12}^2 - \epsilon_2^2}. \quad (\text{C.1})$$

The first step is to identify the set of relevant regions. For this, one may note that the interaction term $\frac{\phi_{12}^2}{\phi_{12}^2 - \epsilon_2^2}$ is different from one only in the *region* in which ϕ_{12} is of order ϵ_2 . We thus define two separate integration regions as follows

$$D^S : |\phi_{12}| \sim \mathcal{O}(\epsilon_2), \quad \text{and} \quad D^L : |\phi_{12}| \gg \epsilon_2, \quad (\text{C.2})$$

where **S** and **L** denote short and long distance respectively, and the whole integration domain is given as $D = D^S \cup D^L$.

The next step is to Taylor expand the integrand in each region. To do so it is convenient to perform the change of variables $\phi_0 := \frac{\phi_1 + \phi_2}{2}$, $\tilde{\phi} := \phi_1 - \phi_2$. In the region D_S one Taylor-expands the Q functions for small¹⁰ $\tilde{\phi}$, and in region D_L one expands the interaction term $\frac{\tilde{\phi}^2}{\tilde{\phi}^2 - \epsilon_2^2}$. Accordingly we introduce the Taylor expansion operators as follows

$$\mathbb{T}^S[I] = \sum_{k=0}^{\infty} \mathbb{T}_k^S[I] = \sum_{k=0}^{\infty} \mathcal{Q}_k(\phi_0) \tilde{\phi}^{2k} \frac{\tilde{\phi}^2}{\tilde{\phi}^2 - \epsilon_2^2}, \quad (\text{C.3})$$

$$\mathbb{T}^L[I] = \sum_{l=0}^{\infty} \mathbb{T}_l^L[I] = \sum_{l=0}^{\infty} Q(\phi_1)Q(\phi_2) \left(\frac{\epsilon_2^2}{\tilde{\phi}^2} \right)^l, \quad (\text{C.4})$$

where

$$\mathcal{Q}_k(x) := \frac{1}{2^{2k}} \sum_{m=-k}^k \frac{(-1)^{k+m}}{(k+m)!(k-m)!} Q^{(k-m)}(x) Q^{(k+m)}(x). \quad (\text{C.5})$$

After the Taylor expansion, one extends the integration region to be the full region D . According to the prescription, one has

$$F = F^S + F^L, \quad F^S = \sum_{k=0}^{\infty} \int_D \mathbb{T}_k^S[I], \quad F^L = \sum_{l=0}^{\infty} \int_D \mathbb{T}_l^L[I]. \quad (\text{C.6})$$

The fact that this equality holds is a very non-trivial statement. In particular one may worry about multiple-counting of contributions. One can verify (C.6) directly upon giving a certain prescription for the evaluation of the relevant integrals. Here we provide a formal argument, following closely [87], which can be easily generalized to the k -tuple integrals entering the instanton partition function. An honest decomposition of (C.1) will proceed as follows

$$F = \sum_k \int_{D^S} \mathbb{T}_k^S[I] + \sum_l \int_{D^L} \mathbb{T}_l^L[I] = F^S + F^L - \left(\sum_k \int_{D^L} \mathbb{T}_k^S[I] + \sum_l \int_{D^S} \mathbb{T}_l^L[I] \right). \quad (\text{C.7})$$

¹⁰Note that ϕ_0 will be equal to the poles given by Coulomb parameters $a_\lambda \gg \epsilon_2$.

In order to show that this is equal to (C.6), one has to show that the terms in the bracket give zero. These terms can be written nicely as

$$\sum_k \int_{D^L} \sum_l \mathbb{T}_l^L [\mathbb{T}_k^S[I]] + \sum_l \int_{D^S} \sum_k \mathbb{T}_k^S [\mathbb{T}_l^L[I]] = \sum_{k,l} \int_D \mathbb{T}_{j,l}^{(S,L)}[I], \quad (\text{C.8})$$

where we use the fact that the double expansion of the integrand is independent of the order in which short or long range expansions are applied¹¹, so that

$$\mathbb{T}_l^L \mathbb{T}_k^S = \mathbb{T}_k^S \mathbb{T}_l^L =: \mathbb{T}_{k,l}^{(S,L)}. \quad (\text{C.9})$$

Explicitly, each term in (C.8) is given as

$$\int_D \mathbb{T}_{j,l}^{(S,L)}[I] = \epsilon_2^{2l} \int \frac{d\phi_0}{2\pi i} \mathcal{Q}_k(\phi_0) \frac{d\tilde{\phi}}{2\pi i} (\tilde{\phi}^2)^{k-l}. \quad (\text{C.10})$$

This is indeed zero since

$$\int d\tilde{\phi} (\tilde{\phi}^2)^n = 0, \quad \text{for integer } n. \quad (\text{C.11})$$

One can take (C.11) as the analogue of setting “scaleless integrals” to zero in the case of Feynman integrals. As in that case, this step is essential in order for the method of expansion by regions to work.

The above considerations seem to be easily generalized to the multi-particle cases. In this case one deals with k -folded integrals. It is expected that the double Taylor expansion in different regions, compare to (C.9), will produce some factorized term in the form of $\int d\tilde{\phi} (\tilde{\phi}^2)^n$. This is set to zero as in (C.11).

Before proceeding to some concrete examples, let us make some important remarks. When applied to the calculation of Feynman integrals, the method of expansion by regions presents the characteristic feature of introducing extra divergences in individual integrals. These new divergences must be regularized and should cancel after summing over all regions in order to reproduce the correct result¹². The situation in the case of the integrals entering the instanton partition function is somehow similar. By staring at (C.3)-(C.4) one realizes that upon Taylor expansion, one produces arbitrary integer powers of $\tilde{\phi}^2$. These are positive powers for $\mathbb{T}^S[I]$ and negative for $\mathbb{T}^L[I]$. The integration along the real line would produce higher order UV divergence in the short regions expansion, and higher order IR divergence in the long range expansion. As in the case of Feynman integrals these divergences needs to be regularized. In the examples below the regularization corresponds to a certain contour

¹¹This corresponds to the case of “commuting expansion” in [87].

¹²For example, in the dimensional regularization $D = 4 - 2\epsilon$, this corresponds to higher order $\frac{1}{\epsilon}$ singularities, corresponding to either UV or IR divergence. The new UV and IR divergences cancel with each other from different regions. These divergence may also be understood as from the scaleless integral, where the UV and IR divergences canceled within the scaleless integral itself (see examples in [87]).

integral interpretation of the integrals. This prescription works in all examples we considered. We recall that the leading term in ϵ_2 , the term we are actually interested in, does not need any regularization of this type.

C.1 Checks of expansion by regions

In this appendix we check the method of expansion by regions via explicit residue computations for two cases. For simplicity, we will consider the $U(1)$ case, namely in the following we take

$$Q(x) = \frac{f(x)}{x - a - i0}, \quad f(x) := \frac{1}{x - a + \epsilon + i0}, \quad (\text{C.12})$$

where we introduce $f(x)$ for convenience as it does not have any pole inside the contour of integration.

Example one

We first consider the integral (C.1). It is straightforward to compute it by summing over the residues of two poles: $(\phi_1, \phi_2) = (a, a + \epsilon_2), (a + \epsilon_2, a)$. One has

$$F = f(a)f(a + \epsilon_2) = \sum_{k=0} \frac{\epsilon_2^k f(a)f^{(k)}(a)}{k!}. \quad (\text{C.13})$$

On the other hand, by the method of expansion by regions, the integral is also given as the sum $F_S + F_L$. One can check that indeed this form reproduces the residue result order by order in ϵ_2 expansion. It is interesting that F_S contributes only to the terms of odd power of ϵ_2 , while F_L contributes to the terms of even power.

Let us spend a few more words on the integrals in F_L

$$\int \frac{d\phi_1}{2\pi i} \frac{d\phi_2}{2\pi i} \frac{f(\phi_1)f(\phi_2)}{(\phi_1 - a - i0)(\phi_2 - a - i0)(\phi_1 - \phi_2)^{2k}}. \quad (\text{C.14})$$

The main issue is how to treat the new denominators ϕ_{12}^{2k} in the expansion. We will apply the following prescription. We first do integration for ϕ_1 . There is a simple pole at $\phi_1 = a + i0$. After taking the residue, the remaining integral for ϕ_2 has a $(2k+1)$ th-order pole of $\phi_2 = a$, the finally residue is given as

$$\frac{f(a)f^{(2k)}(a)}{2k!}. \quad (\text{C.15})$$

This prescription works for the next example as well. This provides a strong support for it.

Example two

We consider an integral which appear at $k = 3$ order in (2.20)

$$F = \int \prod_{I=1}^3 \frac{d\phi_I}{2\pi i} Q(\phi_I) \frac{1}{\phi_{12}^2 - \epsilon_2^2} \frac{1}{\phi_{13}^2 - \epsilon_2^2}. \quad (\text{C.16})$$

As in the example above, one can compute the integral directly by residues. As there is no ϕ_{IJ} terms in the numerators, also poles for which different ϕ 's are at the same position contribute. Let us summarize the result of the residue calculation. The poles of (ϕ_1, ϕ_2, ϕ_3) and the corresponding residues are given as follows:

$$(a, a, a) : \quad \frac{f(a)^3}{\epsilon^4}, \quad (\text{C.17})$$

$$(a, a, a + \epsilon_2) + (a, a + \epsilon_2, a) : \quad -\frac{f(a)^2 f(a + \epsilon_2)}{\epsilon^4}, \quad (\text{C.18})$$

$$(a + \epsilon_2, a, a)^{13} : \quad \frac{f(a)^2 f'(a + \epsilon_2)}{4\epsilon^3} - \frac{f(a)^2 f(a + \epsilon_2)}{2\epsilon^4}, \quad (\text{C.19})$$

$$(a, a + \epsilon_2, a + \epsilon_2) : \quad \frac{f(a)f(a + \epsilon_2)^2}{4\epsilon^4}, \quad (\text{C.20})$$

$$(a + \epsilon_2, a + \epsilon_2, a) + (a + \epsilon_2, a, a + \epsilon_2) : \quad \frac{f(a)f(a + \epsilon_2)f(a + 2\epsilon_2)}{4\epsilon^4}. \quad (\text{C.21})$$

One can see that each pole starts at the order $1/\epsilon_2^4$. However, summing over the residues one finds that the first two orders are cancelled, the leading order starts at $1/\epsilon_2^2$, which gives

$$\frac{6f(a)f'(a)^2 + 3f(a)^2 f''(a)}{8\epsilon_2^2} + \frac{3f(a)f'(a)f''(a) + f(a)^2 f^{(3)}(a)}{3\epsilon_2} + \mathcal{O}(\epsilon_2^0). \quad (\text{C.22})$$

While for our applications, see (2.20), we are interested only in the leading contributions (in the normalization used here this is at order $\frac{1}{\epsilon_2^2}$). In the following we will verify that also the next to leading contribution is reproduced by the method of expansion by regions.

Let us compute (C.16) by the method of expansion by regions. According to the discussion in Section 2.2.2, there are five regions to consider in this case, we have

$$F = F^{\{1,2,3\}} + F^{\{1,2\},3} + F^{\{1,3\},2} + F^{\{2,3\},1} + F^{\{1\},\{2\},\{3\}}. \quad (\text{C.23})$$

In the first region, there is only one cluster, and one only Taylor expands the Q functions around the center of cluster coordinate $\bar{\phi}$, where

$$\bar{\phi} := \frac{\phi_1 + \phi_2 + \phi_3}{3}, \quad \tilde{\phi}_I := \phi_I - \phi_0, \quad (\text{C.24})$$

¹³Note in this case, $\phi_1 = a + \epsilon_2$ should be taken as a double pole, due to $\phi_{12} - \epsilon_2$ and $\phi_{13} - \epsilon_2$ in the denominator.

and $\sum_{I=1}^3 \tilde{\phi}_I = 0$. One has

$$F^{\{\{1,2,3\}\}} = \int d\bar{\phi} \prod_{I=1}^3 \frac{d\tilde{\phi}_I}{2\pi i} \delta\left(\frac{\sum_{I=1}^3 \tilde{\phi}_I}{3}\right) \left(Q^3(\bar{\phi}) + 0 \times \phi_I + \mathcal{O}(\tilde{\phi}_I^2)\right) \frac{1}{\tilde{\phi}_{12}^2 - \epsilon_2^2} \frac{1}{\tilde{\phi}_{13}^2 - \epsilon_2^2}. \quad (\text{C.25})$$

The leading term is straightforward to compute. By integrating out $\tilde{\phi}_I$'s one has

$$\frac{1}{4\epsilon_2^2} \int \frac{d\bar{\phi}}{2\pi i} Q(\bar{\phi})^3, \quad (\text{C.26})$$

which after integrating ϕ_0 exactly reproduces the leading contribution in (C.22). The remaining terms in the Taylor expansion will contribute to $\mathcal{O}(\epsilon_2^0)$ order (since $\tilde{\phi}_I^2 \sim \epsilon_2^2$). These will contribute to the $\mathcal{O}(\epsilon_2^0)$ order in (C.22).

In the next relevant region, there are two clusters. Introduce the center coordinate for the two-particle cluster as

$$\bar{\phi}_a := \frac{\phi_1 + \phi_2}{2}, \quad a = \{1, 2\}, \quad \tilde{\phi} := \phi_1 - \phi_2. \quad (\text{C.27})$$

and set the coordinate in the one particle cluster $\bar{\phi}_b = \phi_3$, where $b = \{3\}$. Doing Taylor expansion one has

$$F^{\{\{1,2\},\{3\}\}} = \int \frac{d\bar{\phi}_a}{2\pi i} \frac{d\bar{\phi}_b}{2\pi i} \frac{d\tilde{\phi}}{2\pi i} \frac{Q^2(\bar{\phi}_a)Q(\bar{\phi}_b) + \mathcal{O}(\tilde{\phi}, \epsilon_2^2)}{\bar{\phi}_{ab}^2 (\tilde{\phi}^2 - \epsilon_2^2)} \quad (\text{C.28})$$

where $\mathcal{O}(\tilde{\phi}, \epsilon_2^2)$ corresponds to higher order terms in the Taylor expansion, which are sub-leading in the small ϵ_2 expansion. To compute the leading term in this region, one needs to evaluate the integral

$$\int \frac{dx}{2\pi i} \frac{dy}{2\pi i} \frac{f(x)^2 f(y)}{(x - a - i0)^2 (y - a - i0)(x - y)^2}, \quad (\text{C.29})$$

where, for simplicity, we considered the case corresponding to the $U(1)$ gauge theory. This integral is similar to (C.14). To evaluate it, one can first integrate out x by taking residue at the double pole $x = a + i0$ then do integration for y . One can also first do integration for y and then for x , the result is the same. By summing the leading contributions from both regions¹⁴ $\{\{1,2\},\{3\}\}$ and $\{\{1,3\},\{2\}\}$, one reproduces the next-to-leading order term in (C.22). It is clear that the region $\{\{1\},\{2\},\{3\}\}$ where all particles are well separated, compare to (C.23), contributes $\mathcal{O}(\epsilon_2^0)$ to (C.22).

This example provides a non-trivial check for the correctness of the method of expansion by regions. In particular, it shows that if one is only interested in the leading contribution, the method of expansion by regions turns out to be extremely efficient.

¹⁴The region $\{\{2,3\},\{1\}\}$ in this case contributes to higher order, since there is no $\frac{1}{\phi_{23}^2 - \epsilon_2^2}$ factor in the integral.

D Alternative derivation of (2.43)

In this Appendix we rederive (2.43) by applying the method of expansion by regions before taking the logarithm of the partition function. This provides a further independent check of the result. The validity of the method of expansion by regions is argued in appendix C, here we mainly exploit the combinatorics coming from the labeling of the regions.

The simple fact on which Mayer expansion is based on is

$$\prod_{1 \leq I < J \leq k} (1 + f_{IJ}) = \sum_{g \in \mathcal{G}^{[k]}} \prod_{e(I,J) \in g} f_{IJ} = \sum_{\mathbf{b} \in \mathcal{B}^{[k]}} \prod_{\mathbf{Y} \in \mathbf{b}} \left(\sum_{g \in \mathcal{G}_c^{\mathbf{Y}}} \prod_{e(I,J) \in g} f_{IJ} \right). \quad (\text{D.1})$$

For convenience we summarize the notation used in the text here: $\mathcal{G}^{\mathbf{X}}$ (respectively $\mathcal{G}_c^{\mathbf{X}}$), where \mathbf{X} is a set, denotes the set of graphs (respectively connected graphs) with the elements of \mathbf{X} as vertices. $\mathcal{B}^{\mathbf{X}}$ is the set of grouping of the elements of \mathbf{X} into clusters. We set $[k] = \{1, 2, \dots, k\}$. Finally, $e(I, J) \in g$ means that there is an edge connecting the vertices I and J in the graph g .

The rewriting (D.1) together with basic properties of the exponential implies that only connected graphs contributes to the logarithm of the grand canonical partition function. The expansion (D.1) is not particularly convenient when expanding the integrand in a region, see Section 2.2.2. There is a refined version of (D.1) that we will now present. For any fixed $\mathbf{b} \in \mathcal{B}^{[k]}$ we can rearrange the integrand as

$$\prod_{I=1}^k U_I \prod_{1 \leq I < J \leq k} (1 + f_{IJ}) = \prod_{\mathbf{Y} \in \mathbf{b}} \mathfrak{U}_{\mathbf{Y}} \prod_{\substack{\mathbf{Y}_1, \mathbf{Y}_2 \in \mathbf{b} \\ \mathbf{Y}_1 < \mathbf{Y}_2}} (1 + \mathfrak{f}_{\mathbf{Y}_1 \mathbf{Y}_2}), \quad (\text{D.2})$$

where

$$\mathfrak{U}_{\mathbf{Y}} := \prod_{I \in \mathbf{Y}} U_I \prod_{\substack{I, J \in \mathbf{Y} \\ I < J}} (1 + f_{IJ}), \quad 1 + \mathfrak{f}_{\mathbf{Y}_1 \mathbf{Y}_2} := \prod_{\substack{I_1 \in \mathbf{Y}_1 \\ I_2 \in \mathbf{Y}_2}} (1 + f_{I_1 I_2}). \quad (\text{D.3})$$

The factors $\mathfrak{U}_{\mathbf{Y}}$, $\mathfrak{f}_{\mathbf{Y}_1 \mathbf{Y}_2}$ respectively takes into account ‘‘interactions’’ within each cluster \mathbf{Y} and between pairs of clusters \mathbf{Y}_1 , \mathbf{Y}_2 . Applying the same reasoning as in (D.1) to the right hand side of (D.2) we rewrite

$$\prod_{I=1}^k U_I \prod_{1 \leq I < J \leq k} (1 + f_{IJ}) = \prod_{\mathbf{Y} \in \mathbf{b}} \mathfrak{U}_{\mathbf{Y}} \sum_{\mathbf{b}' \in \mathcal{B}^{[N_{\mathbf{b}}]}} \prod_{\mathbf{Y}' \in \mathbf{b}'} \left(\sum_{g \in \mathcal{G}_c^{\mathbf{Y}'}} \prod_{e(\mathbf{Y}_1, \mathbf{Y}_2) \in g} \mathfrak{f}_{\mathbf{Y}_1 \mathbf{Y}_2} \right). \quad (\text{D.4})$$

The number $N_{\mathbf{b}}$ denotes the number of clusters in the grouping \mathbf{b} . The sum over $\mathcal{B}^{[N_{\mathbf{b}}]}$ in (D.4) is interpreted as a sum over clusters of clusters.

Assuming the validity of the method of expansion by regions we have

$$\mathcal{Z} = \sum_{k=0}^{\infty} \frac{q^k}{k!} \sum_{\mathbf{b} \in \mathcal{B}^{[k]}} \int_k \mathbb{T}^{(\mathbf{b})} [\text{Integrand}_k] , \quad (\text{D.5})$$

where we define Integrand_k as the r.h.s. of (D.4). The symbol $\mathbb{T}^{(\mathbf{b})} [\dots]$ means the following: Taylor expand its argument in the region labeled by \mathbf{b} . Based purely on combinatorics one rewrites

$$\log \mathcal{Z} = \sum_{k=0}^{\infty} \frac{q^k}{k!} \sum_{\mathbf{b} \in \mathcal{B}^{[k]}} \int_k \mathbb{T}^{(\mathbf{b})} \left[\prod_{\mathbf{Y} \in \mathbf{b}} \mathfrak{U}_{\mathbf{Y}} \sum_{g \in \mathcal{G}_c^{[N_{\mathbf{b}}]}} \prod_{e(\mathbf{Y}_1, \mathbf{Y}_2) \in g} \mathfrak{f}_{\mathbf{Y}_1, \mathbf{Y}_2} \right] . \quad (\text{D.6})$$

The explicit Taylor expansion gives

$$\mathbb{T}^{(\mathbf{b})} [\mathfrak{U}_{\mathbf{Y}}] = \left(\frac{Q(\phi_{\mathbf{Y}})}{\epsilon_2} \right)^{|\mathbf{Y}|} \prod_{\substack{I, J \in \mathbf{Y} \\ I < J}} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} + \dots , \quad \phi_{\mathbf{Y}} := \frac{1}{|\mathbf{Y}|} \sum_{I \in \mathbf{Y}} \phi_I , \quad (\text{D.7})$$

$$\mathbb{T}^{(\mathbf{b})} [\mathfrak{f}_{\mathbf{Y}_1, \mathbf{Y}_2}] = \epsilon_2 |\mathbf{Y}_1| |\mathbf{Y}_2| G(\phi_{\mathbf{Y}_1} - \phi_{\mathbf{Y}_2}) + \dots \quad (\text{D.8})$$

From (D.8) one concludes immediatly that only connected *tree graphs* contributes to the sum in (D.6). This produces a factor of $\epsilon_2^{N_{\mathbf{b}}-1}$. Next, for each cluster we separate the integration variables into center of the cluster and distances between particles in the same cluster. The latter contribution facorizes from (D.7) to produce

$$\frac{1}{\epsilon_2^{|\mathbf{Y}|}} \int \prod_{I \in \mathbf{Y}} \frac{d\phi_I}{2\pi i} \delta \left(\frac{1}{|\mathbf{Y}|} \sum_{I \in \mathbf{Y}} \phi_I \right) \prod_{\substack{I, J \in \mathbf{Y} \\ I < J}} \frac{\phi_{IJ}^2}{\phi_{IJ}^2 - \epsilon_2^2} = \frac{1}{\epsilon_2} \frac{1}{2\pi i} \frac{|\mathbf{Y}|!}{|\mathbf{Y}|^2} . \quad (\text{D.9})$$

We have thus reproduced the result (2.41).

References

- [1] N. Beisert, C. Ahn, L. F. Alday, Z. Bajnok, J. M. Drummond, L. Freyhult, N. Gromov and R. A. Janik *et al.*, “Review of AdS/CFT Integrability: An Overview,” Lett. Math. Phys. **99**, 3 (2012) [arXiv:1012.3982 [hep-th]].
- [2] N. Seiberg and E. Witten, “Electric - magnetic duality, monopole condensation, and confinement in $N=2$ supersymmetric Yang-Mills theory,” Nucl. Phys. B **426**, 19 (1994) [Erratum-ibid. B **430**, 485 (1994)] [hep-th/9407087].
- [3] N. Seiberg and E. Witten, “Monopoles, duality and chiral symmetry breaking in $N=2$ supersymmetric QCD,” Nucl. Phys. B **431**, 484 (1994) [hep-th/9408099].

- [4] A. Gorsky, I. Krichever, A. Marshakov, A. Mironov and A. Morozov, “*Integrability and Seiberg-Witten exact solution,*” Phys. Lett. B **355**, 466 (1995) [hep-th/9505035].
- [5] R. Donagi and E. Witten, “*Supersymmetric Yang-Mills theory and integrable systems,*” Nucl. Phys. B **460**, 299 (1996) [hep-th/9510101].
- [6] H. Itoyama and A. Morozov, “*Prepotential and the Seiberg-Witten theory,*” Nucl. Phys. B **491**, 529 (1997) [hep-th/9512161] ; “*Integrability and Seiberg-Witten theory: Curves and periods,*” Nucl. Phys. B **477**, 855 (1996) [hep-th/9511126].
- [7] E. D’Hoker and D. H. Phong, “*Lectures on supersymmetric Yang-Mills theory and integrable systems,*” hep-th/9912271.
- [8] G. W. Moore, N. Nekrasov and S. Shatashvili, “*Integrating over Higgs branches,*” Commun. Math. Phys. **209**, 97 (2000) [hep-th/9712241].
- [9] A. Losev, N. Nekrasov and S. L. Shatashvili, “*Testing Seiberg-Witten solution,*” In *Cargese 1997, Strings, branes and dualities* 359-372 [hep-th/9801061].
- [10] G. W. Moore, N. Nekrasov and S. Shatashvili, “*D particle bound states and generalized instantons,*” Commun. Math. Phys. **209**, 77 (2000) [hep-th/9803265].
- [11] N. A. Nekrasov, “*Seiberg-Witten prepotential from instanton counting,*” Adv. Theor. Math. Phys. **7**, 831 (2004) [hep-th/0206161].
- [12] N. Nekrasov, “*Five dimensional gauge theories and relativistic integrable systems,*” Nucl. Phys. B **531**, 323 (1998) [hep-th/9609219].
- [13] A. Iqbal and A. -K. Kashani-Poor, “*Instanton counting and Chern-Simons theory,*” Adv. Theor. Math. Phys. **7**, 457 (2004) [hep-th/0212279].
- [14] A. Iqbal and A. -K. Kashani-Poor, “ *$SU(N)$ geometries and topological string amplitudes,*” Adv. Theor. Math. Phys. **10**, 1 (2006) [hep-th/0306032].
- [15] T. J. Hollowood, A. Iqbal and C. Vafa, “*Matrix models, geometric engineering and elliptic genera,*” JHEP **0803**, 069 (2008) [hep-th/0310272].
- [16] A. Iqbal, C. Kozcaz and C. Vafa, “*The Refined topological vertex,*” JHEP **0910**, 069 (2009) [hep-th/0701156].
- [17] L. F. Alday, D. Gaiotto and Y. Tachikawa, “*Liouville Correlation Functions from Four-dimensional Gauge Theories,*” Lett. Math. Phys. **91**, 167 (2010) [arXiv:0906.3219 [hep-th]].
- [18] N. A. Nekrasov and S. L. Shatashvili, “*Quantization of Integrable Systems and Four Dimensional Gauge Theories,*” arXiv:0908.4052 [hep-th].

- [19] C. -N. Yang and C. P. Yang, “*Thermodynamics of one-dimensional system of bosons with repulsive delta function interaction,*” J. Math. Phys. **10**, 1115 (1969).
- [20] A. B. Zamolodchikov, “*Thermodynamic Bethe Ansatz in Relativistic Models. Scaling Three State Potts and Lee-yang Models,*” Nucl. Phys. B **342**, 695 (1990).
- [21] N. A. Nekrasov and S. L. Shatashvili, “*Supersymmetric vacua and Bethe ansatz,*” Nucl. Phys. Proc. Suppl. **192-193**, 91 (2009) [arXiv:0901.4744 [hep-th]].
- [22] N. A. Nekrasov and S. L. Shatashvili, “*Quantum integrability and supersymmetric vacua,*” Prog. Theor. Phys. Suppl. **177**, 105 (2009) [arXiv:0901.4748 [hep-th]].
- [23] S. Cecotti and C. Vafa, “*Topological antitopological fusion,*” Nucl. Phys. B **367**, 359 (1991).
- [24] K. K. Kozlowski and J. Teschner, “*TBA for the Toda chain,*” arXiv:1006.2906 [math-ph].
- [25] N. Nekrasov and E. Witten, “*The Omega Deformation, Branes, Integrability, and Liouville Theory,*” JHEP **1009**, 092 (2010) [arXiv:1002.0888 [hep-th]].
- [26] N. Nekrasov, A. Rosly and S. Shatashvili, “*Darboux coordinates, Yang-Yang functional, and gauge theory,*” Nucl. Phys. Proc. Suppl. **216**, 69 (2011) [arXiv:1103.3919 [hep-th]].
- [27] A. Mironov and A. Morozov, “*Nekrasov Functions and Exact Bohr-Zommerfeld Integrals,*” JHEP **1004**, 040 (2010) [arXiv:0910.5670 [hep-th]].
- [28] A. Mironov and A. Morozov, “*Nekrasov Functions from Exact BS Periods: The Case of $SU(N)$,*” J. Phys. A **43**, 195401 (2010) [arXiv:0911.2396 [hep-th]].
- [29] M. Piatek, “*Classical conformal blocks from TBA for the elliptic Calogero-Moser system,*” JHEP **1106**, 050 (2011) [arXiv:1102.5403 [hep-th]].
- [30] A. Mironov, A. Morozov, Y. Zenkevich and A. Zotov, “*Spectral Duality in Integrable Systems from AGT Conjecture,*” JETP Lett. **97**, 45 (2013) [Pisma Zh. Eksp. Teor. Fiz. **97**, 49 (2013)] [arXiv:1204.0913 [hep-th]].
- [31] N. Dorey, S. Lee and T. J. Hollowood, “*Quantization of Integrable Systems and a 2d/4d Duality,*” JHEP **1110**, 077 (2011) [arXiv:1103.5726 [hep-th]].
- [32] H. -Y. Chen, N. Dorey, T. J. Hollowood and S. Lee, “*A New 2d/4d Duality via Integrability,*” JHEP **1109**, 040 (2011) [arXiv:1104.3021 [hep-th]].
- [33] H. -Y. Chen, T. J. Hollowood and P. Zhao, “*A 5d/3d duality from relativistic integrable system,*” JHEP **1207**, 139 (2012) [arXiv:1205.4230 [hep-th]].
- [34] L. Bao, E. Pomoni, M. Taki and F. Yagi, “*M5-Branes, Toric Diagrams and Gauge Theory Duality,*” JHEP **1204**, 105 (2012) [arXiv:1112.5228 [hep-th]].

- [35] A. Mironov, A. Morozov, B. Runov, Y. Zenkevich and A. Zotov, “*Spectral Duality Between Heisenberg Chain and Gaudin Model*,” Lett. Math. Phys. **103**, no. 3, 299 (2013) [arXiv:1206.6349 [hep-th]].
- [36] K. Bulycheva, H. -Y. Chen, A. Gorsky and P. Koroteev, “*BPS States in Omega Background and Integrability*,” JHEP **1210**, 116 (2012) [arXiv:1207.0460 [hep-th]].
- [37] H. -Y. Chen, P. -S. Hsin and P. Koroteev, “*On the Integrability of Four Dimensional $N=2$ Gauge Theories in the Omega Background*,” JHEP **1308**, 076 (2013) [arXiv:1305.5614 [hep-th]].
- [38] Y. Luo, M. -C. Tan and J. Yagi, “ *$N=2$ supersymmetric gauge theories and quantum integrable systems*,” arXiv:1310.0827 [hep-th].
- [39] M. Aganagic, M. C. N. Cheng, R. Dijkgraaf, D. Klef and C. Vafa, “*Quantum Geometry of Refined Topological Strings*,” JHEP **1211**, 019 (2012) [arXiv:1105.0630 [hep-th]].
- [40] M. -x. Huang, A. -K. Kashani-Poor and A. Klemm, “*The Ω deformed B-model for rigid $\mathcal{N} = 2$ theories*,” Annales Henri Poincare **14**, 425 (2013) [arXiv:1109.5728 [hep-th]].
- [41] S. Hellerman, D. Orlando and S. Reffert, “*The Omega Deformation From String and M-Theory*,” JHEP **1207**, 061 (2012) [arXiv:1204.4192 [hep-th]].
- [42] M. -x. Huang, “*On Gauge Theory and Topological String in Nekrasov-Shatashvili Limit*,” JHEP **1206**, 152 (2012) [arXiv:1205.3652 [hep-th]].
- [43] I. Antoniadis, I. Florakis, S. Hohenegger, K. S. Narain and A. Zein Assi, “*Worldsheet Realization of the Refined Topological String*,” Nucl. Phys. B **875**, 101 (2013) [arXiv:1302.6993 [hep-th]].
- [44] Y. Hatsuda, M. Marino, S. Moriyama and K. Okuyama, “*Non-perturbative effects and the refined topological string*,” arXiv:1306.1734 [hep-th].
- [45] J. Kallen and M. Marino, “*Instanton effects and quantum spectral curves*,” arXiv:1308.6485 [hep-th].
- [46] L. Bao, V. Mitev, E. Pomoni, M. Taki and F. Yagi, “*Non-Lagrangian Theories from Brane Junctions*,” arXiv:1310.3841 [hep-th].
- [47] Hatsuda, Meneghelli, Pomoni, Teschner, Yang, work in progress.
- [48] J. Mayer and M. G. Mayer, *Statistical Mechanics*, New York 1940
- [49] M. Beneke and V. A. Smirnov, “*Asymptotic expansion of Feynman integrals near threshold*,” Nucl. Phys. B **522**, 321 (1998) [hep-ph/9711391].
- [50] M. Gopfert and G. Mack, “*Iterated Mayer Expansion For Classical Gases At Low Temperatures*,” DESY 81/014.

- [51] A. M. Polyakov “*Microscopic description of critical phenomena,*” JETP **55** (1968), 1026-1038.
- [52] D. C. Brydges, *A short course on cluster expansions*, In: *Critical Phenomena, Random Systems, Gauge Theories*. Osterwalder, K., Stora, R. (eds.), Amsterdam: Elsevier 1986
- [53] D. Gaiotto, “ $N=2$ dualities,” JHEP **1208**, 034 (2012) [arXiv:0904.2715 [hep-th]].
- [54] L. Hollands, C. A. Keller and J. Song, “*Towards a 4d/2d correspondence for Sicilian quivers,*” JHEP **1110**, 100 (2011) [arXiv:1107.0973 [hep-th]].
- [55] L. Bhardwaj and Y. Tachikawa, “*Classification of 4d $N=2$ gauge theories,*” arXiv:1309.5160 [hep-th].
- [56] N. Nekrasov and V. Pestun, “*Seiberg-Witten geometry of four dimensional $N=2$ quiver gauge theories,*” arXiv:1211.2240 [hep-th].
- [57] S. Shadchin, “*Cubic curves from instanton counting,*” JHEP **0603**, 046 (2006) [hep-th/0511132].
- [58] M. Billo, M. Frau, F. Fucito, L. Giaccone, A. Lerda, J. F. Morales and D. R. Pacifici, “*Non-perturbative gauge/gravity correspondence in $N=2$ theories,*” JHEP **1208**, 166 (2012) [arXiv:1206.3914 [hep-th]].
- [59] L. Hollands, C. A. Keller and J. Song, “*From SO/Sp instantons to W -algebra blocks,*” JHEP **1103**, 053 (2011) [arXiv:1012.4468 [hep-th]].
- [60] R. Poghossian, “*Deforming SW curve,*” JHEP **1104**, 033 (2011) [arXiv:1006.4822 [hep-th]].
- [61] F. Fucito, J. F. Morales, D. R. Pacifici and R. Poghossian, “*Gauge theories on Ω -backgrounds from non commutative Seiberg-Witten curves,*” JHEP **1105**, 098 (2011) [arXiv:1103.4495 [hep-th]].
- [62] F. Fucito, J. F. Morales and D. R. Pacifici, “*Deformed Seiberg-Witten Curves for ADE Quivers,*” JHEP **1301**, 091 (2013) [arXiv:1210.3580 [hep-th]].
- [63] J. -E. Bourgine, “*Notes on Mayer Expansions and Matrix Models,*” arXiv:1310.3566 [hep-th].
- [64] E. Frenkel and D. Hernandez, “*Baxter’s Relations and Spectra of Quantum Integrable Models,*” [arXiv:1308.3444 [math-QA]].
- [65] N. Nekrasov, V. Pestun and S. Shatashvili, “*Quantum geometry and quiver gauge theories,*” arXiv:1312.6689 [hep-th].
- [66] C. Meneghelli and J. Teschner, “*Lightcone lattice from the universal R-matrix,*” to appear.

- [67] D. Gaiotto, G. W. Moore and A. Neitzke, “*Four-dimensional wall-crossing via three-dimensional field theory*,” Commun. Math. Phys. **299**, 163 (2010) [arXiv:0807.4723 [hep-th]].
- [68] S. Alexandrov and P. Roche, “*TBA for non-perturbative moduli spaces*,” JHEP **1006**, 066 (2010) [arXiv:1003.3964 [hep-th]].
- [69] L. F. Alday, D. Gaiotto and J. Maldacena, “*Thermodynamic Bubble Ansatz*,” JHEP **1109**, 032 (2011) [arXiv:0911.4708 [hep-th]].
- [70] L. F. Alday, D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, “*An Operator Product Expansion for Polygonal null Wilson Loops*,” JHEP **1104**, 088 (2011) [arXiv:1006.2788 [hep-th]].
- [71] B. Basso, A. Sever and P. Vieira, “*Space-time S-matrix and Flux-tube S-matrix at Finite Coupling*,” Phys. Rev. Lett. **111**, 091602 (2013) [arXiv:1303.1396 [hep-th]].
- [72] B. Basso, A. Sever and P. Vieira, “*Space-time S-matrix and Flux tube S-matrix II. Extracting and Matching Data*,” arXiv:1306.2058 [hep-th].
- [73] B. Basso, talks at IGST 2013 and at the workshop “The Geometry and Physics of Scattering Amplitudes”, December 2013.
- [74] B. Basso, A. Sever, and P. Vieira. to appear.
- [75] S. Shadchin, “*On certain aspects of string theory/gauge theory correspondence*,” hep-th/0502180.
- [76] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld and Y. .I. Manin, “*Construction of Instantons*,” Phys. Lett. A **65**, 185 (1978).
- [77] E. Witten, “*Small instantons in string theory*,” Nucl. Phys. B **460**, 541 (1996) [hep-th/9511030].
- [78] M. R. Douglas, “*Branes within branes*,” In *Cargese 1997, Strings, branes and dualities* 267-275 [hep-th/9512077].
- [79] M. R. Douglas, “*Gauge fields and D-branes*,” J. Geom. Phys. **28**, 255 (1998) [hep-th/9604198].
- [80] N. Nekrasov and S. Shadchin, “*ABCD of instantons*,” Commun. Math. Phys. **252**, 359 (2004) [hep-th/0404225].
- [81] M. Marino and N. Wyllard, “*A Note on instanton counting for $N=2$ gauge theories with classical gauge groups*,” JHEP **0405**, 021 (2004) [hep-th/0404125].
- [82] C. A. Keller, N. Mekareeya, J. Song and Y. Tachikawa, “*The ABCDEFG of Instantons and W-algebras*,” JHEP **1203**, 045 (2012) [arXiv:1111.5624 [hep-th]].

- [83] U. Bruzzo, F. Fucito, J. F. Morales and A. Tanzini, “*Multiinstanton calculus and equivariant cohomology*,” JHEP **0305**, 054 (2003) [hep-th/0211108].
- [84] F. Fucito, J. F. Morales and R. Poghossian, “*Instantons on quivers and orientifolds*,” JHEP **0410**, 037 (2004) [hep-th/0408090].
- [85] A. V. Litvinov, “On spectrum of ILW hierarchy in conformal field theory,” JHEP **1311**, 155 (2013) [arXiv:1307.8094 [hep-th]].
- [86] V. A. Smirnov, “*Applied asymptotic expansions in momenta and masses*,” Springer Tracts Mod. Phys. **177**, 1 (2002).
- [87] B. Jantzen, “*Foundation and generalization of the expansion by regions*,” JHEP **1112**, 076 (2011) [arXiv:1111.2589 [hep-ph]].